



universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit

Nominal compounds in English and Serbian:
a comparison

Verfasser

Danijel Marković

angestrebter akademischer Grad

Magister der Philosophie (Mag.phil.)

Wien, November 2011

Studienkennzahl lt. Studienblatt:

A 343

Studienrichtung lt. Studienblatt:

Diplomstudium, Anglistik und Amerikanistik

Betreuerin:

Ao. Univ.-Prof. Christiane Dalton-Puffer

Table of contents

1. Introduction	1
2. The term <i>word</i>	3
3. General remarks on word-formation process in English and Serbian	6
3.1 The role of motivation in word-formation	6
3.2 Word-formation and compounding	8
3.3 What is a compound? – problems of definition	14
3.3.1 Compounding – between morphology and syntax	15
3.3.2 The criteria of a compound	18
3.3.2.1 Structure	19
3.3.2.2 The notion of head	22
3.3.2.3 Stress	24
3.3.3 Compounds vs. noun phrases	27
3.3.4 The morphological and the syntactic-semantic shape and structure of compounds	28
3.4 Compounding and word-formation in Serbian	36
3.4.1 Word-formation in Serbian	36
3.4.2 Synchronic vs. diachronic approach to word-formation	37
3.4.3 The word-formation pattern and the word-formation type	40
3.4.4 The role of motivation and half-motivation in the processes of word-formation in Serbian	43
3.4.5 The notion of root and base	45
3.4.6 The shortening of bases and language economy	46
3.4.7 The role of the binding vowel	47
4. Types of compounds in English	51
4.1 Endocentric compounds	51
4.1.1 Expansions	51
4.1.2 Synthetic compounds	52

4.1.3 Primary or root compounds.....	56
4.2 Copula compounds.....	62
4.3 Exocentric compounds.....	67
4.4 Meaning relations in English compounds.....	71
 5. Composition in Serbian.....	 75
5.1 Prefixation and compounding in the Serbian literature about word-formation.....	78
5.2 What is a compound in Serbian.....	79
 6. Classification of compounds in Serbian.....	 81
6.1 Morphological and syntactic-semantic properties of N-N compounds in Serbian.....	81
6.2 Endocentric compounds.....	85
6.2.1 Determinative compounds.....	85
6.2.1.1 The type noun base + binding vowel + noun.....	87
6.2.1.2 The type noun base + noun (without binding vowels).....	88
6.2.1.3 Compounds derived by affixes (composite-suffixal word-formation).....	89
6.2.2 Complement (rectional) compounds.....	91
6.2.3 Coordinative (copulative) compounds.....	93
6.3 Exocentric compounds.....	95
 7. Conclusion.....	 98
 8. References.....	 101
 9. Appendix.....	 107
9.1 German summary.....	107
9.2 Curriculum vitae.....	109

1. Introduction

If a language is seen rather as a process than as a state, the fact that new words constantly enter a language is not surprising at all. These new words enter the language by many different ways, such as borrowing (taking words from another language), or by creating new words, for instance, by affixation (creating new words by adding suffixes or prefixes), by derivation without affixation (creating new words by conversion), or by compounding (creating new words by putting two words together). All these processes can be comprised by one, more general term, namely the process of word-formation. Even though it may seem as a straightforward division of different processes in word-formation, it becomes clear soon that many of them overlap or do not fit into any of the categories proposed by linguists, as it is the case with neoclassical compounds (e.g. *radioactivity*, *biochemistry*, etc.). However, it is not only neoclassical compounds that cause problems when it comes to word-formation processes, but there are also other types of compounds whose structure is not clearly definable and thus need a comprehensive discussion. Thus, an attempt to describe and compare different types of compounds will be made, with special focus on the compounds that are composed of two nouns, such as *rattle snake*, *can opener*, etc. Additionally, not only compounds from English are going to be taken into account, but also compounds from a typologically quite different language, namely Serbian. As it will become clear in the course of the paper, there are major similarities in the definition and analysis of the compounds in both languages. However, there is a number of differences that must not be neglected and should be looked at in greater detail.

As the term *word-formation* already says the whole process deals with words. Thus the paper will start with the delimitation and the definition of the term *word*. In this chapter an attempt will be made to clarify and exemplify different approaches to this definition. It has to be pointed out that Plag's (2003) theory will be discussed thoroughly, but also other approaches are going to be mentioned and compared with Plag's approach. The second chapter will deal with the general remarks on word-formation processes in English as well as in Serbian, while special attention will be paid to the definition of the terms *compounding* and *compounds* in both languages. Further more, the most important features of

compounds will be discussed, as well as their morphological and syntactic-semantic shape. Additionally, there will also be made a clear distinction between compounds and noun phrases. Apart from these topics, this chapter will also present a comprehensive discussion about the status of compounding in connection with the domains of morphology and syntax, because some linguists see the process of compounding rather as a part of syntax than as a part of morphology. The next two chapters take a look into the classification of noun compounds in English and Serbian. As it will become clear, most of the types that one can find in English, can also be found in Serbian. Finally, the last chapter serves to sum up the most important differences and similarities between noun compounds in both languages, thus enabling the reader to get a clear and systematic overview of different types of noun compounds in English, as well as in Serbian.

2. The term *word*

A compound is a word. But what exactly is a word? Every time we think about it, we take it for granted that everyone knows what a word is. However, as soon as one tries to look closer at the definition, it will become clear that this is not an easy task at all. As Adams (1973: 7) states, the failure of general linguists to provide a consistent definition of the word across languages “has shown that it can only be defined with respect to a particular language; but it is also evident that a word-like unit is equally central and unmistakable for speakers of very diverse languages.” So, for instance, words in English can be simple, which means that they are composed of one constituent only, like *desk*; or they can be complex, which means that they contain more than one element, such as *blackbird* or *undesirable*. The components of the complex words may be free elements: that means that they can appear independently in other contexts, as those in *blackbird*; or they can be bound elements: prefixes and suffixes, which never appear independently, such as the first and the last element of the word *undesirable*.

Plag (2003: 4-5) says that one might think of the word as a “unit in the writing system, the so called **orthographic word**.” Therefore, Plag proposes that a compound could be defined as a string of letters which is preceded by a blank space and followed by a blank space or a punctuation mark. If we have a look at the example (1), this looks like an easy definition and easily applied.

(1) It is raining.

In this sentence one can count three orthographic words. All of them are preceded by a blank space, two of them are also followed by it and the third is followed by a punctuation mark. However, things are not always straightforward. Let us consider the following example:

(2) Mary’s boyfriend has a low-budget job.

According to Plag, the number of orthographic words one finds in this sentence will depend on a number of assumptions. If one considers apostrophes to be punctuation marks, then *Mary’s* is two words. If not, it is only one. The same is the case with the hyphen. Depending on how we see it, *low-budget* could be seen as one or two orthographic words. The most problematic is the word *girlfriend*. Orthographically, it is without any doubt one word. However, most of us would

agree that it consists of two words. Thus Plag concludes the discussion by saying that purely orthographic criteria are unreliable and that one should consider some other aspects. As he claims, the word could be defined in four different ways: in terms of sound structure (phonologically), in terms of its internal integrity, in terms of meaning (semantically), or in terms of sentence structure.

When it comes to the phonological definition of the term *word*, one could say that a word is a unit of speech surrounded by pauses. However, if we listen to the natural speech, it is very likely that people will not make pauses before or after every single word. Therefore, it would be necessary to change the definition. Maybe it would be more appropriate to say that words are units of speech that are surrounded by potential pauses. (Plag 2003: 5-6)

When we consider the integrity criterion, it could be stated that a word is an indivisible unit into which no intervening material may be inserted. If some modifications of the word are to be made, they must be made at the edges, but never inside the words. So, for instance, plural endings such as *-s* are always added at the end of the word, negative elements such as *un-* or *dis-* always at the beginning of the word. However, if we have a look at the words such as *son-in-law*, it could be stated that the law of integrity is violated, since the plural of this word is not **son-in-laws*, but *sons-in-law*. Even though there are some counterexamples to the integrity criterion, it still helps us to come closer to the definition of what a word is. (Plag 2003: 6-7)

When it comes to the semantic definition of the word, it could be said that one word should express one semantic concept. However, this is not always the case. Plag (7) gives the example of the smell of fresh rain in a forest in the fall. Even though this is one semantic concept, we would not consider *the smell of fresh rain in a forest in the fall* as one word. The same problem is with *the woman who lives next door*. The phrase refers to one particular person and should therefore be considered as something defining a unifying concept. This concept is however expressed by more than one word. What we learn from this is that a word may always express a unified concept, but not every unified concept is expressed by only one word. Thus, it seems that the semantic criterion could not help us very much in the defining the term *word*. (Plag 2003: 7)

In the syntactically oriented criterion words are usually seen as syntactic atoms, or the smallest elements in a sentence. These words belong to certain classes (nouns, verbs, adjectives, etc.) which are also called parts of speech, word classes, or syntactic categories. The position in which a word may occur in a sentence is defined by the syntactic rules of a language. For instance, *the* is an article and there are certain rules that say in which position it may occur (usually before nouns as their modifiers). This is why we can test whether something is a word by checking whether it belongs to such a word class. If the item in question follows the rules for articles, it should be an article, and therefore a word. Plag 2003: 8)

Even though the previous discussion may seem to offer a complete definition of the term 'word', Plag states further that even if we agree that a word is a string, there is the problem of ambiguity. If someone pronounces /bi:/, we would not be sure whether the person is thinking of the verb *to be*, or of the insect *bee*. Even though it seems that one cannot give the only and true definition of the word, Plag gives us a deep insight into it and without any doubt helps to better understand the notion of the term *word*.

3. General remarks on word-formation processes in English and Serbian

As it has been stated in the previous chapter, compounds in English are created through the process of word-formation that is called compounding. Even though the aim of this paper is not to explore all the possibilities for creating new words in English, it is necessary to give a brief overview of the most important word-formation processes. The reason why this is important is the fact that one should be able to differentiate between processes of compounding and derivation. It is not possible to say that there is only one way of describing word-formation processes in English, since different authors approach the problem in different ways. It is generally agreed (Plag 2003; Yule 1996; Marchand 1969) that the basic types are derivation, compounding, acronyms, conversion, blending, and clipping. For the purpose of this thesis, it is necessary to make a delimitation only between derivation and compounding, since, as it will become clear later in this paper they overlap very often.

Derivation or affixation is by far the most common word-formation process in English and the whole process is “accomplished by means of a large number of small ‘bits’ of the English language which are not usually given separate listings in dictionaries.” (Yule 1996: 69) These small bits (such as *-less* in *careless*, *re-* in *re-write*) are called suffixes if they appear after the root (e.g. *-less*) or prefixes if they appear before the root (e.g. *re-*). (Plag 2003: 11) Some authors (Yule 1996: 69) point out that there is also a third type of affixes, namely, infix. This kind morpheme is incorporated inside another word. However, their presence in English limited to a very small number of words (e.g. *Hallebloodylujah!*).

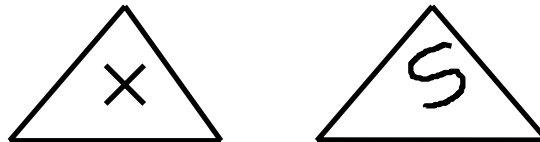
3.1 The role of motivation in word-formation

The word-formation processes exist in all languages and it is a fact that they have a motivated status. In other words, they are analysable sign combinations. According to Saussure¹, a simple linguistic sign is arbitrary, which means that there is no reason why one content should be associated with a certain form. To say this differently, meaning is nothing but a convention that is valid within one speech community. However, it seems that Saussure was radical in some points

¹ For further discussion see Saussure *Course in general linguistics*

since he rejected the idea that in spite of the arbitrariness of the sign, there are some exceptions in the way that the relationship between content and form might be motivated by some external reason, the form of the referent. In those cases, the sign would imitate the referent. Such signs are called icons or iconic signs in semiotics². Classical examples are traffic signs:

Figure 1: Traffic signs



The cross is an iconic representation of two streets crossing and S represents iconically the way the street is turning. (Kastovsky 2008: 12) Similar signs are found today everywhere in the metro stations, airports, or railway stations because in this way the passengers do not have to rely on their linguistic knowledge if they want to find a toilette or to have a cup of coffee. There are some other signs that Saussure also excluded, namely, the sign in the domain of sound imitations. These words, the so-called onomatopoeic words, such as *bang*, *crash*, *ding-dong* can also be treated as motivated by their extralinguistic referent whose properties these lexical items imitate. According to Kastovsky (12-13) linguists assign much greater importance to iconicity than Saussure used to do and they agree that onomatopoeias are extralinguistically motivated, but that this motivation is not complete³.

Now, let us turn to sign combinations, which are the primary domain of word-formation. This is where “the principle of compositionality [...] is central”. (Kastovsky 2008:14) Kastovsky argues that sign-combinations are motivated in relation to their constituents and other formally and semantically parallel signs. Sometimes this parallelism produces one, two or three combinations that are formed by direct analogy, as in:

(3) einsam → zweisam, ?dreisam, *viersam

² For further discussion about semiotics see Peirce *Semiotische Schriften*

³ For further discussion about motivation see Marchand (1969: 397-439)

Some analogies can be extended until they become a productive pattern. In these cases the meaning of the sign combination can be derived on the basis of the meanings of the constituents and some other underlying constructional meaning.

It is this basic computability of meaning which constitutes the compositionality principle, and which is the underlying principle of everything in language (word-formation, morphology and syntax). (Kastovsky 2008: 15)

Kastovsky gives a very striking example of numerals in every language where there is always a subdivision of a minority of numerals that are arbitrary and the great majority of other numerals whose meaning can be derived on the basis of the arbitrary numerals and some other patterns which are motivated. Thus, in English the numerals 1-12 are arbitrary, while the combination 14-19 are analysable signs, consisting of a number and *-teen* which can be seen as an allomorph of *ten*. Number 20, 30, 90, etc. can be analysed in the same manner: *two* + *-ty*, *three* + *-ty*, *nine* + *-ty*, *-ty* meaning 'ten times'.

Thus, it can be stated that one of the most important aspects of word-formation is their relatively motivated status. By this motivation, Kastovsky (2008: 15) assumes that there is a certain logical correlation between the meaning and the form and in this way the meaning can be derived from the form. In this way the unsystematicity that is a part of the arbitrary sign is restricted. However, it should be mentioned that different languages show different tendencies towards the motivated words. So, for instance, Kastovsky (16) claims that German exhibits a much more stronger tendency than English. Let us have a look at the following examples:

- (4) a. Mund : mündlich
b. mouth: oral

To sum up, as it can be seen from the previous examples, motivation is based on formal grounds, because there is a certain morphological and formal transparency. This is best seen in the examples of signs that are composed of more than one sign and this is exactly what is relevant for this thesis, namely, word-formation and compounds.

3.2 Word-formation and compounding

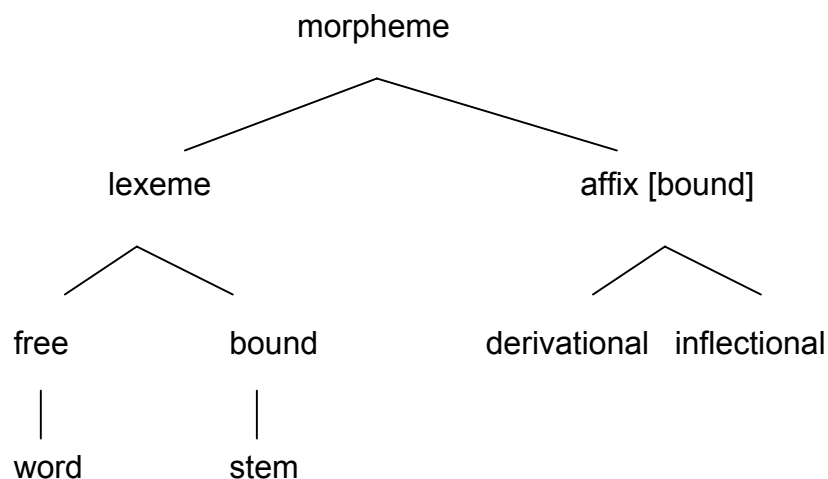
Word-formation is a part of the science of language that studies the patterns for creating new lexical units or words. (Marchand 1969:2) Therefore, word-formation

can only be concerned with composites or words that are a combination of full signs. In other words, composites are words that are made of morphemes, the smallest meaningful units, which in this case are the smallest linguistic signs. Another feature that defines a composite is the relation of the morphemes within the composite on the determinant/determinatum basis. Or in other words, those are signs based on a significate/significant relation.⁴ In contrast to primary lexemes “which are arbitrary and have to be learned in isolation, syntagmas [composites] are relatively motivated.” (Perkles 2008: 4) Once it is agreed that we are dealing with complex words, different examples should be examined. Let us have a look at the following composites:

- (5) a. highway, blackbird, writing table
 b. re-write, undo, inability
 c. neighbourhood, flawless, novelist

The examples represent different possibilities of combining morphemes in order to create a complex lexeme, which in the same time show the various possibilities of creating them: compounding and affixation.

(6) Kastovsky (2008: 44)



The first thing one has to do in order to describe what kind of word-formation process is taking place, it is necessary to look at the structure of the complex lexeme. Let us start with the examples in (5). In this case the composite (syntagma, complex lexeme) is made of two different free morphemes, or in other words, two different lexical items, as in *highway* (*high* + *way*), *blackbird* (*black* + *bird*), *writing table* (*writing* + *table*). The last example, *writing table* differs from the

⁴ For further discussion about significate/significant relationship see Saussure

previous two in a way that it consists of three, and not of two morphemes (*writ-ing* + *table*). However, they are joined by the process of compounding which includes binding two lexemes together.

The examples in b. and c. are slightly different. They fit the rules of the syntagma proposed by Marchand, however they are different from compounds in the way that they are made of a free lexeme and a bound morpheme:⁵ *re-write* (*re* + *write*), *undo* (*un* + *do*), *inability* (*in* + *ability*), *neighbourhood* (*neighbour* + *hood*), *flawless* (*flaw* + *less*), *novelist* (*novel* + *ist*). Depending on the position of the bound morpheme, linguists usually speak of prefixed or suffixed syntagmas, or simply of prefixation and suffixation. It is important to mention that in some cases the affix (prefix, suffix, infix) does not have to be overtly expressed, e.g. *water*. This is an instance of what Marchand (1969: 359) calls 'derivation by a zero morpheme'. In this concrete example, the noun *water* is used to derive a verb, through adding of the zero morpheme at the end of the noun:

(7) *water* + Ø → *to water*

Even though we cannot see it, it has a particular grammatical function and meaning since it converts the noun *water* into a verb. Some linguists (Plag 2003, Bauer 1983) also call it conversion, since the word changes its class without actually changing its form. Thus, it can be concluded that compounding is joining two lexemes together, while under affixation we understand joining a free lexeme and a bound morpheme. (Perkles 2008: 5) This definition is unfortunately very loose and does not hold for many other examples. Let us have a closer look at the following composites taken from Marchand (1969: 356-358):

- (8) a. *manlike*, *apelike*, *beastlike*, *snakelike*
b. *praiseworthy*, *seaworthy*, *trustworthy*
c. *clockwise*, *sidewise*, *coastwise*

The main question here is which part of these complex words can we see as independent morphemes? For words such as *man*, *ape*, *clock* or *sea* it is easy to give an answer. However, what about the second part of the words? What we have in these cases is what Marchand (1969: 356) describes as words that stand "midway between full words and suffixes". Why does he claim so? To answer this

⁵ For further discussion about bound morphemes see Plag (2003: 10)

The situation in examples in b. is quite similar to those in a. Namely, *worthy* and *like* are both adjectives historically. And again Marchand stresses their semi-suffixal character by the fact that “adjectives which they form can be prefixed by un- [unpraiseworthy]. Compound adjectives cannot be prefixed.” The examples in c. are not any less controversial. Complex lexemes built with *wise* have the meaning ‘in the form, manner, or like of’. What is problematic about the morpheme is that there is a word *wise* in English whose status as an independent word is unquestionable. Phrases like *a wise man* or *a wise thought* are possible and in this case the meaning of the word *wise* is slightly different than those in complex lexemes. Clearly, it means ‘clever or intelligent’. Marchand (1969: 358) also points out that today *wise* is “being used less and less as an independent word”, except in some phrases that are used rarely, e.g. *in any wise*, *in no wise*, or *in gentle wise*. (Perkles 2008: 6)⁶

(9) a. fights of mock - literary dialogue
b. a funky, regional blues - type version (Renouf & Baayen 1998)
c. handful full
d. Celtic-Type type (Dalton-Puffer & Plag 2001)

⁶ For detailed discussion about semi-suffixes see Marchand 1969

than free lexemes. However, the shift from derivation to compound may happen too. Fernández-Domínguez (2008: 70-71) states further that this shift can affect both prefixes and suffixes. Consider the following examples:

- (10) a. Sexism ism
 b. Hungrish ish
 c. Photograph photo
 d. Television tele (Bauer 2005c (cited in Frenandez): 101)

Marchand (1969: 326) calls these elements *affixoids* or *semi-suffixes* and as it has been stated earlier Marchand claims that those are instances of forms that stand midway between full words and suffixes. Dalton-Puffer & Plag (2001: 243) argue however that

some previously proposed analyses need to be revised. In particular, it was argued that there are good arguments for treating these formations as either compounds (in the case of *-type*), or as affixes (in the case of *-ful* and *-wise*). Labels such as ‘semi-suffix’ are theoretically undesirable and do not provide additional insights into the nature of complex words.

As can be seen, it is the fact that the meaning of the affixoids is narrower when they occur within a compound when they are free forms (e.g. *like* and *worthy*). They are usually seen “as a case of grammaticalization where semantic but no formal change has taken place (hence their resemblance to compounds).” Fernández-Domínguez (2008: 71) However, Booij (2005a (cited in Fernandez): 114-117) states that not all forms of this type should be considered affixoids because it is not possible to use them for novel compound-creation, as in (11):

- (11) a. playwright
 b. shipwright
 c. horemonger
 d. moneymonger
 e. scandalmonger

There is one more type of compounds that should be discussed in a detail, namely the compounds that have been formed by the use of analogy. These words are unusual in so far as that they are “derived without an existing word-formation rule, but formed on the basis of a single (or very few) model words.[...] The process by which these words come into being is called analogy[.]” (Plag 2003: 37) Let us have the following examples from Plag:

- (12) a. a : b :: c : d
 b. eye : eyewitness :: ear : earwitness

- c. ham : hamburger :: cheese : cheeseburger
- d. sea : sea-sick :: air : air sick

The essence of this process lies in the formula in a. The relation between items *a* and *b* is the same as the relation between the two others. Thus the relation between *eye* and *eyewitness* is the same as the relation between *ear* and *earwitness*. In the same manner, the relation that holds between *ham* and *hamburger*, holds in the same way between *cheese* and *cheeseburger*. It is also worth mentioning that the situation with the *burger*-series of compounds needs to be discussed in more detail. Kastovsky (2008: 20) states that this type of compounds has become productive in an unusual way and manner. The starting point here was *hamburger steak*, which a special kind of dish which was prepared from ground or minced meat and shaped in a flat, round manner. Somehow, the dish was associated with the immigrants from Germany, thus bringing one to the conclusion that *ham* in the sense of ‘cured pork’ has nothing to do with hamburgers. People then shortened the term *hamburger steak* to *hamburger* and started experimenting with different spices and types of meat, in this way producing different kinds of hamburgers. The term was once again shortened to *burger* and adopted the meaning “flat, round dish from ground ingredients, served between a roll.” (Kastovsky 2008: 20) This meaning became productive and resulted in the compounds that we have today.⁷

Until now only compounds with two simple elements have been discussed (*dust + bin*, *arm + chair*). However, a very productive group of compounds is the one that consists of one simple and one complex element (*letter + writer*, *dining + room*), or of two or more complex elements (*drugstore + owner*). Even though we describe them as compounds, there are some other processes of word-formation involved in the creation of this kind of compounds. Let us have a look at the structure of these words:

- (13) a. truck driver: truck + er (V + er) > truck + driver (N + N)
 b. dining room: dine + ing (V + ing) > dining + room (N + N)
 c. drugstore owner: own + er (V + er) > drug + store (N + N) > drugstore + owner (N + N)

In the example a. the composite cannot be regarded as an accidental binding of the morphemes *letter + writ + er*, but rather as a formation that has undergone a

⁷ See Kastovsky (2008: 20)

straightforward process. The starting point for the formation is the verb *drive*, to which we have to add the suffix *-er* in order to create a noun. After this has been done it is possible to combine it with the noun *truck* in order to create the compound *truck driver*. In the same manner, the example in b. does not include only the process of compounding. The first step is to create a noun from the verb *dine*. This is done by adding the suffix *-ing* at the base, and now it is possible to combine it with the morpheme *room*, by which we get the compound *dining room*. The example in c. is slightly different from the previous two in so far as it involves more than two steps. Namely, it is necessary to add the suffix *-er* to the verb *own* in order to create the agent, but before it is possible to combine it with the second part of the compound, we need to join the words *drug* and *store* together. Thus, within one compound we find the processes of suffixation and compounding within a compound.

3.3. What is a compound? – problems of definition

When the problem with the definition of the term *word* is solved, one can turn to the definition of the compound. As it has already been stated at the beginning of this thesis, the definition of a compound does not follow a straightforward line. Many linguists have dealt with this topic and have come up with different ideas about compounds, and therefore the following section is going to be devoted to a short historical overview of the discussions relating to compounds. The purpose of this overview is not only to present different theories, but also to compare them. Once all important definitions are presented, it might be possible to choose one that could be applied to the compounds dealt with in this thesis.

The coining of new words proceeds by way of combining linguistic elements on the basis of a determinant/determinatum relationship is called syntagma. When two or more words are combined into a morphological unit on the basis just stated, we speak of a compound. (Marchand 1969: 11)

This is the way Hans Marchand defines compounds. The reason why one should start with his definition is that “he is considered, especially in Europe, to be the ‘father’ of the modern word-formation theory”. (Štekauer 2000: 29) However, even though he was one of the first to give such a detailed overview of the word-formation in English, one should not forget that it has almost been half a century since the book was published. Therefore, it is logical to expect that some

of his ideas have been modified or changed. For that reason, after examining Marchand's theory, it is necessary to explore the theories of other influential morphologists.

3.3.1 Compounding – between morphology and syntax

The assumption in this paper is that compounding is a process of word-formation, and not a syntactic process, as some linguists claim. Before going into any deep discussions about compounds and the process of word-formation, it is necessary to say what the process of word-formation means. Marchand (1974a: 171) states that “[w]ord-formation is that part of grammar which studies the patterns on a which language forms new lexical units” and it is to be assumed that a derivative (new lexical unit) is a syntagma that consists of a determinant and determinatum. Another important concept in this context is the principle of ‘motivated lexical item’. This means that meanings of the words can be derived from the meaning of their constituents and some kind of a morphosemantic pattern. This morphosemantic pattern

represents the constructional meaning of the word-formation in question and it can in its most general and abstract form be reduced to a binary relation of the following kind: one constituent, the determinatum (or head), is determined, i.e. specified, by the other constituent, the determinant (or specifier, modifier)[.] (Kastovsky 2008: 20)

Thus, it does not matter whether we have a compound (e.g. *armchair*), a suffixal derivative (e.g. *child-hood*), or a prefixal derivative (e.g. *re-write*), all of them can be analysed on the same determinant/determinatum principle. Such a construction based on the determinant/determinatum relationship is called syntagma by Marchand (1969: 11) Both parts of the previously named words are semantic morphemes, “signs based on a signifiant/signifié relation.” As Marchand (1974a: 173) argues further

[i]n the event full of compounds, the syntagma is opposable to either morphemic element. [...] Prefixal and suffixal derivatives must be opposable to their unprefixal and unsuffixal bases [...] and to other derivatives containing the same dependent morpheme[...]

In other words, the compound *armchair* has to be opposable to the noun *chair*, the prefixed derivative *re-write* needs to be opposable to *write*, and the suffixed derivative *childhood* to *child*. In some cases the derivative element may be absent

(e.g. *convert* V vs. *convert* N) and in this case we speak of zero morpheme. However, since derivation with zero morpheme is not the main issue of this paper, this problem will not be discussed any further.⁸

Let us now come back to compounds. Compounds are those words that are in the English language qualified as 'complex words'. These complex words are products of the process that is called in linguistics 'derivational morphology' or simply 'word-formation'. Thus, if we agree that complex lexemes are products of word-formation and that a compound is a complex lexeme, the logical conclusion that follows from these premises is that a compound is word-formation. Kastovsky (1982: 155) argues that products of word-formation can any time become a part of our lexicon and that it is not necessary to create these words every time we want to use them, since they are stored and ready-made in our lexicon and we can refer to them any time we want. This is easily supported by the fact that words such as *armchair*, *rainbow*, *dustbin*, or *headache* are stored in our mind as such and we do not have to form them from two separate words when we want to use them. This feature of word-formation products is what Kastovsky (1982: 155) calls 'Inventarcharakter'.

The term 'word-formation' is used to define the process of creating new words in a language and it is important to stress that there are different ways of creating words. Therefore it is necessary to ask what the process in word-formation is that forms compounds. The only logical answer to this question would be 'compounding'. Thus, the process of compounding operates in the domain of word-formation. The understating of this process enables us to say "what a compound is, [to] recognise it in a given context or speech situation, or even [to] use the rules of its production as a model for building new ones". (Perkles 2008: 4)

At the beginning of the chapter it has been stated that some scholars see compounding as a syntactic process and not as a morphological one. This claim is extremely controversial and if it turned out to be true, everything that has been stated about the theory of language would have to be re-written. Thus, it is necessary to have a deeper look into this proposal. Plag (2003: 159) states that some proponents of this theory argue that a very productive subclass of compounds, namely noun-noun compounds, results from a syntactic rule that

⁸ For further discussion about zero-derivation see Marchand (1969: 359-389)

states that in a certain noun phrase not only adjectives, but also nouns can be the modifiers of the following noun. Let us look at the following examples:

- (14) a. NP → article {adjective, noun} noun
 b. the long marathon
 c. the Vienna marathon

The brackets in a. show that the modifier of the noun can be either an adjective or a noun. As it can be seen from the examples in b. and c. the head noun is in one case modified by an adjective and in another by a noun. Looking at these instances, one could easily conclude that proponents of the syntactic theory are right. So far so good. The problem arises when it comes to the analysis of the stress in the compounds/phrases. Plag notices that this rule does not explain why adjective-noun combinations usually have their stress on the noun (Lieberman & Sproat 1992: 134), and therefore have the “flavour of phrases” (Plag 2003: 159), while noun-noun combinations are usually stressed on the first noun (Lieberman & Sproat 1992: 134) and therefore have the “flavour of words”. (Plag 2003: 159)⁹ According to this, one could say that the rule stated in the previous example does not exist. It is too early to say this since in certain constructions nouns can be analysed as phrasal premodifiers of other nouns. Consider the following data:

- (15) the Vienna market, a three-syllable word, the two-year period

One could argue that these are examples of compounds. However, if we try to insert adjectives between the words in (15) it is possible, and therefore cannot be seen as compounds, but as nominal phrases. The same test cannot be applied to the real compounds, as it can be seen in (16b):

- (16) a. The Vienna financial market, a three-syllable prosodic word, a two-year probationary period
 b. rain forest vs. *rain tropical forest

This leads us to the stress criterion that is going to be discussed in connection with compounds. The examples in a. have in common that they are stressed on the right-most element, while the examples in b. have leftward stress. Thus, this may be taken as an indication that instances in a. really are phrases. If we assume that they are phrases, then there is nothing unusual about the fact that an adjective may be inserted between the modifier and the head. To be more specific, it seems

⁹ For the exact percentage of left-stressed and right-stressed compounds see Lieberman, Mark & Richard, Sproat 1992

that stress patterns and the syntactic behaviour together strongly argue in favour of a phrasal analysis of the specific constructions. This would of course mean that all compounds are phrasal, or at least, that all compounds with the final stress are phrasal. Plag (2003: 160) argues the opposite. He says that there are only some “restricted classes of nouns whose members are allowed to act as syntactic modifiers of nouns.” According to him, those two classes are nouns indicating location and nouns incorporating a numeral. When it comes to stress, Marchand (1969: 21) stated:

For a combination to be a compound one condition has to be fulfilled: the compound must be morphologically isolated from a parallel syntactic group. [...] Blackbird has the morphophonemic stress pattern of a compound, black market has not, despite its phrasal meaning; the latter therefore is a syntactic group, morphologically speaking. Stress is a criterion here.

Later however he says “[a]s for the criterion of stress, we shall see that it holds for certain types only.” (Marchand 1969: 29)

To sum up, in order to be sure whether a certain construction underlies the rules of syntax or word-formation, it is necessary to decide what kind of a construction it is. In the previous example, we classified some instances as compounds and some as phrases. In those cases where we speak of compounds, we also speak of word-formation, and in those where we speak of phrases, we speak of syntax. This seems as a rather simple solution. However, the greatest problem actually is how to differentiate between noun phrases and compounds. Thus, the following two sections are going to deal with this topic.

3.3.2 The criteria of a compound

As far as the straightforward definition of compounds is concerned, it can be stated at the very beginning of this chapter that the aim of this thesis is not to find the only right definition of compounds, but to find one that would be most suitable. There are several features that need to be taken into consideration, when the description of the compounds is concerned. However, before discussing the most important features of compounds, it is necessary to mention that spelling is not going to be seen as one of the criteria for compounds, since it is very unreliable. Compounds are sometimes written as two separated words, sometimes as two hyphenated words and sometimes as one word.

3.3.2.1 Structure

Selkirk (1982: 13-14) claims that the structure of compounds can be explained easily by looking at the possibilities of combining different word classes in order to create a compound. She states that compounds in English are a type of words made up of two constituents, where each of the constituents may belong to the categories noun, adjective, verb, or preposition. The compound itself may belong to the categories noun, verb, or adjective. More specifically, a compound noun may consist of a noun, adjective, preposition, or a verb on the left and a noun on the right, a compound adjective may consist of a noun, adjective, or a preposition on the left and an adjective on the right side, and a verb may consist of a preposition followed by a verb, as represented in (17). According to Marchand (1969: 80), Jespersen (1909-1949, vol. IV: 143-145), Bauer (2003: 202-207) N + N is the most productive type of compounds, and thus this is going to be the main task of this paper, namely, to have a close look at the structure of this type.

(17) a. Nouns				
N + N		A + N	P + N	V + N
apron string		high school	overdose	swearword
armchair		well-wisher	uprising	scrubwoman
achoolteacher			inland	rattlesnake
b. Adjectives				
N + A		A + A	P + A	V + A
headstrong		ice cold	overwide	none
honey-sweet		hardworking	underprivileged	
heartbroken		widespread		
c. Verbs				
N + V		A + V	P + V	V + V
none		none	outlive	none
			overdo	
			offset	
			underfeed	

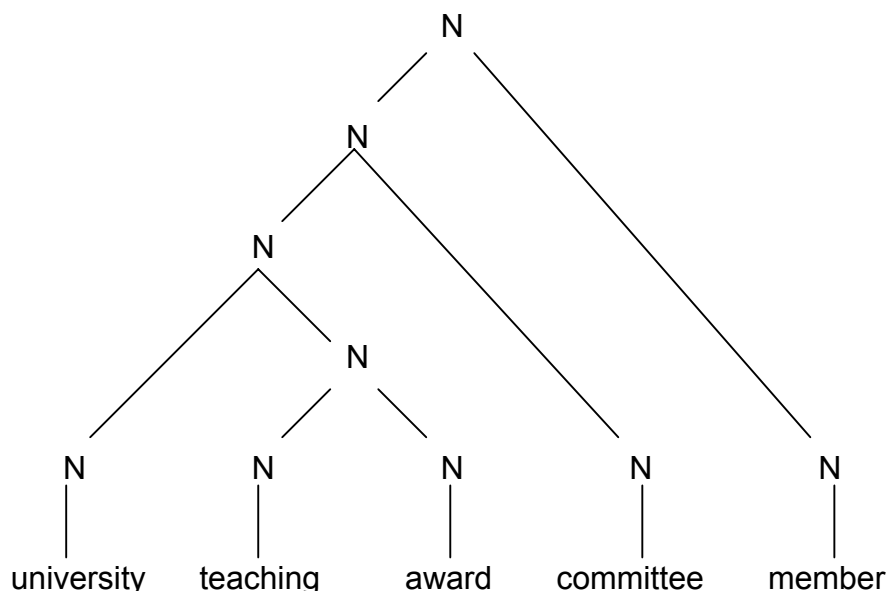
She also states that more complex structures are possible because of the recursive structure of compounds. The notion of compounding is sometimes called composition and as already stated it is to be expected that compounding has something to do with joining two words together in order to form a new word. This statement has two crucial assumptions: the first is that a compound consists of two (and not more than two) elements, and the second is that these elements are words. At first sight, this seems to be a rather straightforward definition. However, there are some examples that do not quite fit the definition and need to be

discussed in a greater detail. Let us look at the following example from Plag (2003: 133): *university teaching award committee member*. As it can be seen, the statement that a compound consists of only two words seems to be wrong. This is actually partly true, since it is generally possible to analyse complex words as hierarchical structures involving binary subelements. According to this rule, the analysis of the compound would look as follows:

(18) [[[university[teaching award]]committee]member]

As the bracketing shows, the compound can be divided into strictly binary compounds as its constituents. The innermost constituent [*teaching award*] (an award for teaching) is made up of [*teaching*] and [*award*], the next larger constituent [*university teaching award*] (the teaching award of the university) is made up of [*university*] and [*teaching award*], the constituent [*university teaching award committee*] (the committee responsible for the university teaching award) is made up of [*university teaching award*] and [*committee*], and finally the [*university teaching award committee member*] (member of the committee responsible for the university teaching award) is made up of [*university teaching award committee*] and [*member*]. This can also be represented by the tree diagram (Figure 2):

Figure 2



means that they consist of two elements, but not of two words (as already discussed, one element can be more than one word).

It is interesting what Plag (2003: 134) mentions further about compounds (at least for noun-noun compounds). It is always possible to add a new word to a compound in order to create a new compound. Thus, it is possible to create a compound like this one: *university teaching award committee member training* (training for the members of the university teaching award committee). This property is called 'recursivity' and it is known from the analysis of sentence structure. For instance, the grammar of the English language allows us to use subordinate clauses recursively, by putting a new clause inside each new clause. For example, *Mary said that Harry knew that John believed that Sarah thought...* and so on. As we have seen, it is possible to do the same with the compounds. However, one should not use very long compounds too often since, on the one side it is difficult for the speaker to produce such constructions, and on the other it is difficult for the listener to comprehend them.

After we have concluded that compounds can be analysed as binary structures, it is necessary to answer the question which units (elements) can be used to make a compound. Let us consider the following examples from Plag (2003: 134):

- (19) a. astrophysics, biochemistry, photoionize
b. parks commissioner, teeth marks, system analyst
c. pipe-and-slipper husband, off-the-rack dress, over-the-fence gossip
d. *husband pipe-and-slipper, *dress off-the-fence, gossip over the fence

The first examples are interesting because they consist of elements that are not attested as independent words in the English language (*astro-*, *bio-*). Plag (74) defines them as bound roots.¹⁰ Thus, it would be suitable to redefine compounding as the combination of roots, and not words. However in the second example we see that the first element is not a root but a plural form, which is a (grammatical) word. To make the whole thing worse, in the third example we see that even larger units (syntactic phrases) can occur as the elements of the compound. Phrases usually occur on the left side of the compound. If they are on the right side, it is either impossible (**husband pipe-and-slipper*), or they are not interpreted as

¹⁰ Plag (2003: 74) argues that if one analyses the elements of the word *biology* in the following way: *bio-* being a prefix and *-logy* being a suffix, then we would have a paradox that a word consists of a prefix and a suffix. For that reason Plag proposes to call these elements bound roots.

compounds (*gossip over the fence*). Taking into account all these statements, it could be said that a compound “is a word that consists of two elements, the first of which is either a root, a word or a phrase, the second of which is either a root or a word.” (Plag 2003: 135)

3.3.2.2 The notion of head

The vast majority of compounds in English is interpreted in that way that the left-hand member somehow modifies the right-hand member. So, a *stone wall* is a kind of a wall, a *film society* is a kind of a society. Therefore, one can say that English compounds exhibit what is called a modifier-head structure. The term head

is generally used to refer to the most important unit in complex linguistic structures. In compounds it is the head which is modified by the other member of the compound. Semantically, this means that the set of entities denoted by the compound [...] is a subset of the entities denoted by the head. (Plag 2003: 135)

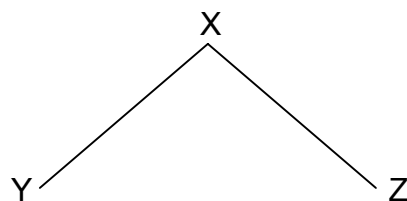
The head usually occurs at the right-hand side of the compound (the so called right-hand head rule) and this is a very important systematic property of compounds. The compound as whole receives most of its semantic properties from its head. So, for instance, if the head of the compound is a verb, the compound will be a verb (e.g. *deep-fry*), if the head is a noun, the compound will be a noun (e.g. *beer bottle*), if the head has feminine gender, the compound will have the feminine gender (e.g. *head waitress*). Another very important feature of the compound head is that the plural marking occurs on the head and not on the non-head. Thus, the plural of the compound *book cover* is not **books cover*, but *book covers*. Plag (2003: 137) concludes that considering the constituency and the headedness of compounds, one can formalize the structure of compounds as follows:

- (20) a. [XY]y
b. X = {root, word, phrase}
Y = {root, word}
y=grammatical properties inherited from Y

The differences between phrases and compounds will be discussed later in the paper, as the differences among them are best seen when the stress of compound is concerned.

As already stated, a vast majority of compounds in English are endocentric¹¹ (Selkirk 1982: 19), which means that they have a head and the head is on the right. According to Selkirk (19) there are few types that do not fit into this pattern, but they are clearly endocentric. Those compounds belong to the verb-particle group (e.g. *grow up*, *sit down*, *step up*). And they obviously have their head on the left side. There is also a small group of the so called endocentric compounds, which do not have head at all. The head of compounds plays the crucial role in the description of the distribution of the diacritic features that are related to both inflectional and derivational morphology. In this context it is necessary to introduce the term ‘percolation’, a feature that ensures that a constituent and its head have the same feature complex. Selkirk defines percolation in the following way: “[i]f a constituent α is the head of a constituent β , α and β are associated with an identical set of features (syntactic and diacritic).” (19) In other words, this would mean that a compound as a whole “must be marked for many the same features as the head of the compounds.” (Bauer 2003: 177)

Figure 3



Thus, given a tree such as in Figure 3, where X is an endocentric compound, and Y and Z are its constituents, and where Z is the head of the construction, it means that X must carry many of the grammatical features as its right-hand constituent, in this case it would be Z. Bauer (178) also points out that features from both of the constituents are available to the compound as a whole. In some cases, these features will be in agreement, and in some cases they can be in conflict. For instance, in the compound *seabird* the constituents will agree in that way that they are both nouns. However, from the semantic point of view they are in conflict because sea is inanimate, while bird is animate. If there is a conflict, the whole compound receives the features of the head element. Thus, the compound *seabird* is an animate noun. If there is not any conflict, the compound is marked by the features of the both constituents.

¹¹ See chapter about endocentric compounds in English

3.3.2.3 Stress

One of the first things all linguists start with when they define compounds is to look at the stress patterns of the word. Just as with all the other features, it is not possible to say if this criterion is the one that can be used for defining a compound. “[W]e shall see that it holds for certain types only”, argues Marchand (1969: 21). On the whole, all linguists agree that compounds tend to have a stress pattern that is different from that of phrases. While phrases tend to be stressed phrase-finally, which means on the last word, compounds tend to be stressed on the first element. Plag (2003: 137) argues that the systematic difference is captured in “the nuclear stress rule (‘phrasal stress is on the last word of the phrase’) and the [...] compound stress rule (‘stress is on the left-hand member of the compound’).”

- (21) a. Noun phrase:
 [this new hóuse]
 b. Nominal compound:
 [installátion guide]

It is very often the case that the stress pattern is the only thing that makes the distinction between a compound and a noun phrase; in other words the stress can lead to minimal pairs where it seems that the stress pattern is the only thing that distinguishes between a noun and a compound (the most prominent syllable is marked by an acute accent on the vowel). The compounds on the left are characterized by having the main stress on the left constituent, and the weak stress on their second member. The weak stress (or middle stress as called by Marchand (1969: 28)) is marked ` . This pattern ´/ is also called ‘forestress’, ‘unity stress’, or simply ‘compound stress’. In contrast, phrases on the right have two strong stresses (also termed heavy stress by Marchand (28)) and this pattern is termed as ‘level stress’.

(22)

Noun compound

bláckboard

‘a board to write on’

óperating instructions

Noun phrase

a black bóard

‘a board that is black’

operating instrúctions

‘instructions for operating something’

‘instructions that are operating’

The compound stress rule can account for most of the compounds. However, there are always exceptions to the rule as the examples in (23) illustrate (the most prominent syllable is marked by an acute accent on the vowel):

(23)

geologist-astrónomer

apple píe

Michigan hóspital

Madison Ávenue

summer níght

aluminium fóil

may flówers

silk tíe

As it can be seen, these examples do not fit to the rule that has been stated previously, namely, that all compounds have their stress on the first part. If this is not a compound, what is it? If it is a compound, how can we account for it? Plag (2003: 138) recommends an alternative approach. Thus, he argues that these examples are not “idiosyncratic but that they are more or less systematic exceptions of the compound stress rule.” (138) The rightward prominence is restricted to only well defined types of meaning relationships. If one starts with the compound *geologist-astrónomer*, it is special for that reason that both elements of the compound refer to one entity. One person is a geologist and an astronomer at the same time. Such compounds are called copulative or dvandva compounds and they are going to be discussed in greater detail in the chapter about types of nominal compounds. It is important to mention that this subclass of compounds systematically shows the rightward stress and therefore it can be stated that it is a systematic exception to the rule. Plag argues further that other meaning relationships that are systematically accompanied by the rightward stress are temporal and locative meanings (e.g. *a summer night*, *Michigan hospital*), or causative meaning, which is usually paraphrased by ‘made of’ (e.g. *aluminium foil*). When it comes to the example *Madison Ávenue*, it seems that all compounds that have the word *avenue* as a part of them tend to have the rightward stress,

while all compounds that have the word *street* as their right-hand member tend to have the leftward stress (e.g. *Óxford Street*).

Now, let us explore the other hypothesis, which says that only words with the leftward stress are compounds, which implies that examples above are not compounds. If they are not compounds, what are they? The easiest solution to account for this phenomenon is to say that they are phrases. However, this would create other problems. Let us start with the *avenue* and *street* examples. Both words have the same internal structure (noun – noun), they show the same meaning relationship between their constituents, and they are both right-headed. The only difference between them is that one of them has leftward and the other rightward stress. Hence, it seems impossible to claim that *Madison Avenue* is a phrase and that *Óxford Street* is a compound just because they have different stress patterns. Thus, taking into account all the previous statements, according to Plag (2003: 139) compounds in English generally have leftward stress. The counterexamples are usually systematic exceptions that only appear in certain types of semantic relation between the compound elements.

If one considers the rules stated above as a general rule for all compounds (with some systematic exceptions), another problem arises, namely, the stress of the compounds that have more than two elements. The following examples could be taken into consideration:

- (24) a. *máil* delivery service vs. mail *delí*very service
b. *stú*dent feedback system vs. student *fé*edback system
c. *gó*vernment revenue policy vs. government *ré*venue policy

As it can be seen from the examples, the stress pattern can indicate a certain type of interpretation. (Plag 2003: 140). Thus, a *máil delivery service* is a *service* concerned with *máil delivery*, while a *mail delívery service* is a *delívery service* concerned with *mail*. The same interpretation can be applied to the other examples. Therefore, a *stúdent feedback system* is a *system* concerned with *stúdent feedback*, whereas a *student féedback system* may be interpreted a *féedback system* that has something to do with *students*. Similarly, the *góvernment revenue policy* is the *policy* that is concerned with the *góvernment revenue*, while the *government révenue policy* is a *révenue policy* as implemented

by the *government*. The two different interpretations can be illustrated by using brackets:

- (25) a. [[máil delivery] service] vs. [mail [delivery service]]
b. [[stúdent feedback] system] vs. [student [féedback system]]
c. [[góvernment revenue] policy] vs. [government [révenue policy]]

Looking at these examples Plag (2003: 140) concludes that in the three pairs given the most prominent syllable is always placed on the left-hand member of the compound within the compound, but never on the member that is not a compound itself. Liberman and Prince stated that in a compound that has a [XY] structure, Y will receive the strongest stress, if, and only if, it is a compound itself. In other words, a compound [XY] will have a left-hand stress if Y is not a compound itself. If Y is a compound, the rule is applied again to Y.

3.3.3 Compounds vs. noun phrases

Even though we have established some rules for recognizing a compound, it can happen that one cannot differentiate between, for example, a compound and a noun phrase. As we know, a noun may be premodified by another noun, an adjective or particle, or a nominalization. The resulting sequence may be a free phrase or a compound in which the premodifying element has lost its independence. So, the question is, how should one decide which of the sequences *small talk* and *wet day* is a free phrase and which one is a compound? In order to be able to do this, one should apply tests to find out whether the first elements are separable from the heads. (Adams 1973: 57) Can the adjective be modified by an adverb? It is possible to say *very wet day*, but it is not possible to say *very small talk*. Can we use the comparative form of the adjective? Again, it is possible to say *wetter day*, but not *smaller talk*. Can the adjective be used in the predicative function in a sentence with the head noun as its subject? It is possible to say *The day is wet*, but it is not possible *The talk is small*. Thus, it is not difficult to conclude that *small talk* is a compound, while *wet day* is a free phrase.

If the first part of the sequence is a noun, one should apply tests on which of the elements, the first or the second, can actually be modified by an adjective? Let us look at the following list with examples:

- (26) a. wooden door way, outer door knob, rare book case
b. narrow door way, brass door knob, mahogany book case

In the first set of compounds, the adjective is an appropriate modifier of the second element, while in the second set of examples, the adjective is an appropriate modifier of the third element. Since the first set of compounds sounds odd, we can assume that *door way*, *door knob* and *book case* are compounds, since the first element cannot be modified independently of its head.

3.3.4 The morphological and syntactic-semantic shape and structure of compounds

The first task that one has to do before one starts to analyse compounds is to give their morphological description. This means that one has to give a description of the morphemic elements that form the compound. This can be done in the following way: *armchair* = *arm* N + *chair* N, *craftsman* = *craft* N + *s* + *man* N. It is presupposed that a speaker has a clear knowledge of the content of the morphemic constituents and in that way all the ambiguities (that can arise from homophony for instance) are excluded. Morphemes are signs and without knowing their semantic meaning we would not be able to distinguish and describe compounds. Let us have a look at the following examples: *corn belt* and *safety belt*. In order to describe the compounds it is necessary to differentiate between the words *field*. Both of them basically denote the same sign. However, *belt* in *corn belt* denotes a 'belt-like area', while in *safety belt* we have the meaning of 'belt as considered a device or instrument'. (Marchand 1989: 54)

The next step is to state the immediate constituents of compounds and give their morphological description. According to Marchand (1989: 54) all morphologic compounds are based on the same syntagmatic pattern 'determinatum determined by determinant'. The structure of compounds in English is determinant/determinatum, meaning that the determinant precedes the determinatum (e.g. in *steamboat*, *steam* narrows down or determines the meaning of *boat*; in *wind mill*, *wind* determines the meaning of *mill*). On this restricting basis of the determinant/determinatum it is easy to predict the meaning of the compound. "The semantic content of the constituents in conjunctions with grammatical deep structures [...] and morphologic structure assign a compound a meaning that only deliberate joking can sometimes construe differently or give a

new interpretation.” (Marchand 1989: 54) In order to understand the determinant/determinatum principle, let us look at the following examples:

- (27) a. wind mill
 b. play writer
 c. football player
 d. holiday car sightseeing trip

The compound in a. has a simple determinatum, which means that it is a primary lexeme. However, in the examples b. and c. the second constituent is a suffixed derivative and in these cases the determinata are complex. As it can be seen in the example c. the determinant can also be a complex word. When it comes to the morphological importance of the determinatum, it is necessary to stress that it is the dominant element only because of the reason that it establishes the word class of combination. For instance, *wind mill* is a noun compound because the determinatum (*mill*) is a noun. The same applies to all examples in (27). The other important morphological property of the determinatum is the fact the determinatum reflects the way how a compound is inflected. Therefore, the plural of *wind mill* must be *wind mills*, and not **winds mill*. The formation in d. is for several reasons a special case and needs to be discussed in a greater detail. As it can be seen, the rule that a compound can be broken down into two units, can also apply here. However, the determinant in this compound is itself a complex unit. The question is if this is possible. Marchand (1974a: 199-200) argues that in a compound only the determinant can be a compound. If it turns out that the determinatum is a compound, then the whole unit should be seen as a syntactic group. Plag (2003: 135) on the other hand gives a very loose definition of the structure of compound and states that it consists of two elements, the first of which is either a root, a word, or a phrase, the second of which is either a root or word. He does not give any further details about the number of elements that these words may consist of. Carstairs-McCarthy (2002: 76-77) suggested that those compounds that have the structure as the ones in d. can be called compounds and that a. have a compound for the determinant, as well as for the determinatum. This view arises from the fact that we need to identify a compound as one whole, morphologically complex unit. (Perkles 2008: 11) Marchand (1969: 21-22) proposed that for a complex unit to be treated as a compound it is necessary that it is possible to morphologically isolate it from a parallel syntactic group. What is meant by morphological isolation is the

fact that a compound has to be stressed differently from similar syntactic constructions (e.g. *bláckboard* ‘a board to write on’ vs. *a black bóard* ‘a board which is black’).¹² According to this definition, different accents decide whether a unit should be treated as a compound or as a syntactic group. Again, Marchand and Plag have different approaches when it comes to accents. Marchand (1969: 28) differs between the heavy stress (marked ´), middle stress (marked `), and weak stress which is traditionally called absence of stress, while Plag (2003: 137-141) distinguishes only between the strong stress (marked ´) and the weak stress (which is the same as Marchand’s middle stress). Usually, compounds have the heavy (primary) stress on the left constituent and the middle stress on the second one. A syntactic group has the primary stress on both of its constituents. In these cases it is stress that is responsible for the recognition of a compound and the only morphophonemic difference between compounds and similar syntactic phrases.

Now, let us come back to the example in d. In this example we find two compounds as constituents. In those combinations the stress is not on the compounded determinant, but on the determinatum. As it is a compound itself, the stress that is usually on the first constituent is moved to the left member of the determinatum. Thus, the stress in the compound *holiday car sightseeing trip* would be on *sight*. However, as it has already been discussed in the chapter about stress, there are numerous examples that do not fit this rule. Some of the following examples exemplify these borderline cases: *ápple cake*, or *Mádison Street* vs. the syntactic constructions *apple píe*, or *Madison Ávenue*. *Ápple cake* and *apple píe* are both cakes made of apples and *Mádison Street* and *Madison Ávenue* both refer to the same concept and are named after their discoverer or inventor. The problem how to classify these instances has not been solved until today and in most books they are referred to as exceptions to the rule.¹³

However, in order to be able to interpret and describe compounds, morphological description only is not enough. At this point it is necessary to touch upon the syntactic-semantic structure of compounds. Again, the starting point for the discussion will be Marchand. He assumes that all composites are syntagmas, which in other words would mean that they are grammatical entities. This fact

¹² For further discussion about stress see chapter 3.3.2.3

¹³ For further discussion see Plag (2003: 137-141) and Marchand (1969: 24-29)

implies that they must be explainable from an underlying sentence whose syntactic relation they actually mirror. Marchand (1969: 55) calls this grammatical relationship grammatical deep structure. Let us have a look at the following examples: *truck driver*, *steamboat*. The compound *truck driver* is explainable from '(we) drive the truck', while the compound *steamboat* is derived from 'steam operates the boat'. Marchand also points out that the underlying sentences should not contain composite forms (e.g. *maker*, *driver*, *happiness*) as they are themselves nominalizations of sentences. As it has been stated, nominal composites are derived from a sentence and this underlying sentence will take different nominal forms depending on what we want to express. According to Marchand (1969: 32) one grammatical part of the sentence is taken to be known: the subject, the object, the predicate, the predicate complement, or the adverbial complement. This is the part of the compound that becomes the determinatum, while its syntagmatic complement in the sentence, the part that gives further information with regard to the determinatum "invariably becomes the determinant". (32) These selectional patterns of information are called 'types of reference'. Let us take the sentence 'we drive trucks' as an example. Each of the grammatical parts is to be found in the sentence. Therefore the subject type (S) would construct the compound *truck driver*, the predication type (Pr) *truck driving*, and the object type (O) *driving truck*. As it can be seen, the name of each type of reference indicates which part of the sentence is made the determinatum of the compound.

When it comes to the subject type, the information about the subject may not only refer to its activity, but it may also include the complement of this activity (direct object or adverbial complement). The subject may be expressed overtly by a substantive or a suffix, or it may be expressed by a zero morpheme. Let us have a look at the following examples:

- (28) a. dancing girl
 b. chimney sweep
 c. car dealer

In the example a. the compound is derived from the sentence 'the girl dances'. The subject is a full noun, subject and activity are expressed and the verb may be transitive or intransitive. The example b. has the underlying structure 'he sweeps chimneys'. However, the subject is not overtly expressed, but by a zero morpheme

(*chimney + sweep + Ø*). *Car dealer* is derived from ‘he deals in cars’ and in this case the subject is marked by the suffix *-er*. There are a few things that should be said about the word order of this type of compounds. The word order of sentences is the opposite of that in compounds. Sentences show the “progressive dm/dt order [...] whereas morphologic composites as a rule follow the reverse dt/dm type.” (Marchand 1969: 34) To exemplify, the sentence ‘he deals in cars’ has the structure subject – predicate – prepositional object, while the compound *car dealer* has the opposite structure of that in the underlying sentence.

In the object type information can only contain a statement about the relation Predicate – Object, where the object is expressed by a substantive, by a suffix, or by a zero morpheme. It should also be mentioned that the direct object of the active sentence is the subject of the passive sentence. Let us analyse the following examples:

- (29) a. *mincemeat*
 b. *government employee*
 c. *convert*

Mincemeat is derived from ‘the meat has been minced’ (P –O); *government employee* has the underlying structure ‘has been employed (by the government)’ (P –O (sf)); *convert* is derived from ‘has been converted’ (P – O (Ø)). (Marchand 1969: 35)

When it comes to the predication type, nothing is known but the fact that something, a kind of an activity is going on. What we have to do then is to describe what kind of an activity it is. This question of what kind the activity is, demands nouns of different predication types. This activity may be connected with the subject (event, activity caused by the subject) or with the object (event, activity is directed towards the object) of the sentence. According to Marchand the clear subject-oriented cases are derived from intransitive verbs (e.g. *arrive*, *live*) while the object oriented cases are derived from transitive verbs (e.g. *arrest*, *defeat*). Marchand also points out that the predication type is partly tied up with syntactic conditions. Whenever a deverbal substantive is followed by an *of*-group, the subject-oriented type arises. In this case the subject has an active sense (e.g. ‘his release’ and ‘his release of the prisoners’).

Now, let us have a look at the following examples:

- (30) a. sunrise
b. alcohol intake
c. well-being
d. play going

In a. the subject is included in *sunrise* from the 'sun rises', and this type affects only intransitive verbs (e.g. *waterfall* from 'water falls'). (Marchand 1969: 36) In *alcohol intake* we find the object included in the compound. The compound itself is derived from 'they take in alcohol'. The example c. has the predicate complement included, since *well-being* is derived from 'their well-being is...', and the adverbial complement is included in the example d., as *play going* can be derived from 'go to plays'.

What is characteristic about the adverbial complement type is the fact that we receive information about "the circumstances or activity". (Marchand 1969: 37) Marchand gives a list of several different instances:

- a. place where the action is performed: *writing table* from 'we write at a table'; *drive-in restaurant* from 'we drive in the restaurant'
- b. instrument with which the action is performed: *washing machine* from 'we wash with the machine'
- c. time when the action is performed: *closing time* from 'they close at that time'

This division of compounds into the types of reference by Marchand is not the universal and only one. Apart from this distinction, compounds can also be divided according to their semantic head, i.e. whether they have one, whether it is on the left or on the right side, or what kind of morpheme it is. Before we start discussing the classification of compounds, it is necessary to say something about the semantic properties of compounds.

A compound has two immediate constituents, namely the determinant and the determinatum. When it comes to the semantics, it has to be stated that the determinatum plays a more important role than the determinant, because of the fact that it decides the lexical class of the whole compound. Let us take a couple of examples: *armchair*, *dance hall*, *truck driver*. In all these cases it is the right constituent (determinatum) that is in charge of the lexical class of the compound. Thus, *armchair* is not just a chair, but a chair where you can rest your arms upon; *dance hall* is a hall built for dancing, and *truck driver* is not any kind of driver, but

the one that drives trucks. In connection with this, it would be useful to mention the concept of hyponymy. Kastovsky (2008: 89) defines hyponymy as “the relationship between a more general and one or several more specific terms”. What is meant here is the fact that the determinant somehow modifies the meaning of the determinatum, which at the same time means that it “narrows down its applicability”. (89) Thus, the constituent *arm* narrows down and modifies the meaning of *chair*, just as *truck* gives a more precise description of *driver*. In other words, a compound functions as a subordinate term of a more general one, which means that a compound is nothing but the hyponym of its head. Kastovsky (89) calls this more general term archilexeme. Thus, in the example *armchair*, the archilexeme is CHAIR. However, as far as the semantic properties of compounds are concerned, the syntactic relations that can exist between the immediate constituents should be pointed out. (Perkles 2008: 16) Let us have a look at the following examples:

(31)	N	+	(V-er)		dt	dm
	letter	+	writ-er	=	letter	writer
	tape	+	record-er	=	tape	recorder

Both of these compounds, *letter writer* and *tape recorder*, have the same morphological structure; they both consist of a simple noun as determinant and a derivative as determinatum. Even though they have the same morphological structure, they cannot be analysed in the same way semantically. Thus, *letter writer* is someone who writes letters, while *tape recorder* is something for recording on tapes. In the first case the compound denotes a person (profession), while in the second case the compound denotes a thing (something). It is the suffix *-er* that is responsible for the difference in analysing the determinatum. In a. it produces an agent (someone doing something), while in the second it creates an instrument (something is used for something else). Thus in *letter writer* the relation between the determinant and the determinatum is that of object and person, while in *tape recorder* it is one of place and instrument. These different semantic relations are closely connected to syntactic ones. So, *letter writer* could be paraphrased as ‘someone who (S) writes (P) letters (O)’, which means that it would have the O-P-S structure. For this reason the terms ‘word-formation pattern’ and ‘word-formation type’ are introduced. Word-formation pattern has to do with the morphological possibility to form a compound (e.g. N + (V-er)), while word-

formation type refers to a particular syntactic-semantic relation between the constituents (e.g. O-P-S). (Perkles 2008: 17)

As it can be seen, in analysing compounds there are several factors that need to be taken into consideration. It is not enough just to look at the morphological structure, just as it is wrong to not to consider the syntactic-semantic relation between the constituents of a compound. They all correlate with each other and none of it should be left out. Kastovsky (1982: 214-215) relying on Marchand gives a short summary of the most important points that one needs to consider when analysing compounds. Marchand (1969: 53) calls this “patterns for description of composites”. Consider the example *wall painter*, Kastovsky’s five points will be presented:

- (32) a. morphological shape: wall paint-er = N + (V-er)
b. morphological structure: wall/paint-er
c. syntactic-grammatical structure (underlying/deep structure): ‘someone (S) paints (P) walls (O)’
d. type of reference: subject type
e. meaning on the morphological level (semantics): ‘a person whose profession is to paint walls’

The first step is the morphological shape of the compound, because it helps us to find out more about the type and status of its constituents. The next step would be to determine the immediate constituents of the compound, and in this way it is possible to say which morphemes are modifiers, and which one is the head. Since Marchand (1969: 55) claims that all compounds are analysable from underlying sentences, the logical step would be to find the paraphrase of the compound (the underlying sentence). After this has been done, it is possible to establish the type of reference by relying on the deep structure. Finally, the lexical meaning of the compound is derived on the basis of the relation between the constituents. When it comes to the compounds discussed in this paper, it has to be mentioned that it will be necessary to deal with these points selectively, which means that for some types only certain descriptions would be necessary and maybe the only possible solution. Kastovsky (1982: 215) says that “[e]ine solche Analyse bildet den Ausgangspunkt für die synthetisch-generative Darstellung der Wortbildung.“ It will be seen later, especially in the chapter about exocentric compounds, that these formal rules cannot be applied, or applied only partially.

After the most important properties of compounds have been discussed, it is possible to ask the question: what kinds of compounds are there in English and Serbian? Since the subject of this paper is nominal compounds (only those that consist of two nouns), and for length reasons, other types of compounds in English and Serbian are not going to be discussed in a greater detail.

3.4 Compounding and word-formation in Serbian

3.4.1 Word-formation in Serbian

For a long time word-formation in Serbian has dealt only with suffixation, which means that other types of creating new words were not taken into account, or if it was the case, then in an extremely small amount. (Klajn 2002: 5) For instances, Maretić (1899: 292) in the introductory chapter about general word-formation in Croatian and Serbian states that “[k]ako u svim drugim srodnim jezicima tako se i u hrvatskome ili srpskom riječi sastoje u dvome: u korijenu i nastavku (jednome) ili u nastavcima (ako ih je više od jednoga)” and then he continues the discussion only about roots and suffixes. The possibility that a word might consist of two roots is not considered at all. In the same manner Stevanović (1964: 401) in his grammar also writes only about the ‘root’ or ‘the general part’ and the ‘ending’ or the ‘suffix’. Only towards the end of the chapter does he mention composition in one sentence stating that one part of word-formation is also composition, however, without giving any examples. After the introductory section, a section about compounds follows, in which prefixed words are also included, and then comes the most comprehensive chapter of the grammar, namely the part about suffixation, which he names derivated words. One of the most influential morphologists Belić (1949: 2) does however recognize that there are simple, derivated and composite words in Serbian. According to him, simple words are those that consist only of root, composite are those made of two words, and derivated are those where one can distinguish between the root (or stem) and the ending. Even though Stevanović and Belić were seen as very influential and respected linguists, the problem with their analysis of compounds lies in the fact that they considered compounding as joining two words together and compounds as a kind of lexicalised syntagma, and thus according to them composition was

something that is a part of syntax, while the only 'real' word-formation is derivation, i.e. suffixation.

Today, in modern linguistics, the general opinion is accepted that there are four different and completely equal types of creating new words in Serbian, namely prefixation, suffixation, composition, and conversion. (Klajn 2002: 6) The fact that they are considered to be equal does not mean that they all deserve the same amount of attention, since they are not equally productive, and thus it is to be expected that the largest amount of literature has been written about suffixation because it is by far the most productive way of producing new words in Serbian, while conversion creates only a small number of new words. Interestingly, none of the authors deals with coinage, blending or acronyms as possible ways of enlarging the vocabulary.¹⁴ Nonetheless, it has to be mentioned that Klajn was the first to give a comprehensive overview of word-formation processes in Serbian that included a large chapter devoted to compounding (Klajn 2002: 15-173) and a chapter (even though a small one) that deals with conversion (Klajn 2003: 379-389).

3.4.2 Synchronic vs. diachronic approach to word-formation

The general trends in linguistics in Serbian have been marked by the fact that there is a strong tendency towards moving to the descriptive from the historical description of grammar. (Klajn 2002: 7) Some of the early linguists, such as Maretić, were not satisfied with the fact that the root is the part of the word that cannot be changed, but he was looking for the root according to the etymology. Thus, this approach excludes the presence of simple words and the number of suffixes rises to the extreme. For instance, for him the noun *mjesec* ('moon') is composed of the root *mjes-* and the suffix *-ec*, *stado* is derived from *sta-* and *-do*¹⁵, etc. The problem with Maretić's analysis is that he tries to find the etymological root in words in which the etymological root is actually unknown. Belić was aware of this problem, however he was not ready to give up the diachronic approach. So, for instance, in his book *Savremeni srpskohrvatski književni jezik*¹⁶ (1986) he writes about the 'real derivatives', words where the root and the suffix were clearly

¹⁴ For further discussion about other types of word-formation processes in English see Yule (1996: 64-70)

¹⁵ Suffixes *-ec* and *-do* do not exist today (Klajn 2003: 397-401)

¹⁶ *The contemporary Serbo-Croatian literary language*

distinguished, and what he calls 'lexically unproductive words', such as words that end in *-va* (*osnova*, *mrkva*, *breskva*).¹⁷ For most of time Stevanović (1964) tries to stick to the synchronic approach, but he is not successful in this all the time. For example, he tries to explain that the noun *sin* ('son') as a derivative from the root *su-* (it used to mean 'be born') and the suffix *-n*. (410) As the suffix *-n* does not exist today, the question is if one can see this a derivative. The first one who definitively broke with the synchronic approach was Babić in *Tvorba riječi* ('Word-formation') from 1986. However, his method was an extreme one and he only concentrated on the formal analysis of words, which included splitting words into their morphemes. The problem is that he analyses morphemes with foreign origins in the same manner as the morphemes that originate from the Slavonic languages. For instance in the chapter about pure compounds (319-327) he states that *fotoćelija* ('photocell') is a pure noun compound, without even mentioning that the first part of the compound has Greek origins. The latest attempt to describe word-formation processes in Serbian was conducted by Klajn in 2002 and 2003. As he states at the very beginning, his approach is a synchronic one, but where necessary one has to look into the etymology of words. (Klajn 2002: 11) In this paper, this approach will be taken into account. This means that all the compounds will be explained from the synchronic point of view. Where it is unclear how the word is structured, one has to look at the etymology of the word in order to be able to analyse it. Thus, the theories of all of the authors mentioned above cannot be applied to every single word, but rather an attempt will be made to make a selective choice as to which theory can be applied best to which word.

As Klajn (2002: 10) argues, the synchronic approach should not be understood as a static one, but as

onaj iseček dijahronijskog prikaza do koga možemo doći na osnovu jezičkog osećanja, bez pribegavanja istorijskoj fonetici, bez pozivanja na ranije jezičke faze ili na mrtve i hipotetičke jezike.[that fragment of the diachronic view that can be reached on the basis of our language feeling, without having to look at the historical phonetics, without having to look at the early language phases or at dead or hypothetical languages].

This language feeling is of course a category that is not easy to comprehend. However, it is still the only reliable norm, and one has to prefer it to the formalist

¹⁷ The suffix *-va* does not exist in the modern Serbian language (Klajn 2003: 397-401)

approach. Klajn (10) argues that this language feeling is assumed. For instance, considering the language feeling of an average Serbian speaker, it is very probable that he or she does not know the etymology of the word *opijum* ('opium'). On the one hand they are probably still aware of the fact that the word has a similar form in other languages, on the other it is known that Serbian verbs do not build derivatives by the suffix *-ijum*; thus for him or her it does not sound convincing that the word *opijum* is a derivative of *opiti* ('become alcoholised'). Etymology can be very helpful in cases where it is not enough to look at the synchronic features when it comes to the categorization of a certain word-formation. For example, *praskozorje* ('dawn') is derived from the syntagma 'vreme kad zora praska' ('the time when the dawn thunders'). The problem with this compound is the fact that it is not known whether the first part of it (*prask-*) is of nominal or verbal origin. It can be analysed either as 'vreme kad zora puca' ('the time when the dawn thunders') or as 'prasak zore' ('the thunder of the dawn'). In cases like this, it would be necessary to take etymology into account. Clearly, it should not be the only criterion, and one should not consider words that have lost its motivation as products of word-formation (e.g. *sin*¹⁸ should not be seen as a derivative because its motivation has been lost). From that point of view, Klajn (2002: 11) and Babić (1986: 15) agree that certain words, such as *riznica* ('treasury'), are not a product of word-formation, because it does not any more mean 'mjesto gdje se čuvaju rize' ('a place where someone keeps coins'), but rather 'a place where someone keeps precious things'. However, on the same page Babić states that a word that is not a product of word-formation is the synonym for unmotivated word. Klajn (11) does not agree with it and exemplifies it with the word *srdžba* ('anger'). For Babić this is not a product of word-formation because the word *srčiti se* (today *srditi se*) ('become angry') does not exist in the same form today and thus *srdžba* is to be excluded from the word-formation processes. However, Klajn claims that these words are without any doubt in a relationship with *srditi se*. In this case, one can say that the root somehow changed or that the suffix has an unusual form, but it cannot be denied that it is a product of word-formation. Further on, Babić (1986: 19) claims that certain names of towns, such as *Karlovac* or *Kraljevica* are 'etymologically transparent words'

¹⁸ See p. 38 for more about *sin*

and then he states that their meaning today cannot be understood as 'Karlo's town' and 'Karlo's settlement' because what they mean today is 'the town at the mouth of Korana and Kupa' and 'the town on the entrance to the Bakarski bay' respectively. Klajn (2002: 11) claims that these 'etymologically transparent words' are nothing but products of word-formation processes. *Karlovac* cannot be analysed in the same way, as for instance *Zagreb*, because every speaker of Serbian or Croatian feels that there must be a certain relationship between a person whose name used to be Karlo and the town *Karlovac*.

Finally, it can be concluded that when it comes to the discussion about synchrony and diachrony, one should always be careful not to go to the extremes. Probably the best way to escape this problem is to use motivation as the most helpful tool to analyse complex words. On the one hand, it is not useful today to claim that words that used to be complex (e.g. *stado*) are still complex today. On the other, one should not deny the relationship between two words where this relationship is obvious, just because one of the words for some reason slightly changed or modified its form. Moreover, as Klajn (2002: 12) argues, motivation does not have to cover the whole word, because there are some instances of the so called 'half-motivation'. According to him, an affix can lose its meaning, but it is still affix because it is added to the root in order to form a new word, and thus as long as there is only one part of the word that undoubtedly connects it to the other, that word must be seen as a complex one and thus fits the scope of word-formation. (Klajn 2002: 7-12)

3.4.3 The word-formation pattern and the word-formation type

When it comes to word-formation processes it is, according to Barić (1980: 12-13), unnecessary to distinguish between the word-formation pattern ('*tvorbeni uzorak*') and the word-formation type ('*tvorbeni tip*'), since they have been used as synonyms by many linguists. She claims that "[t]vorbeni je uzorak ključni pojam u tvorbenom sustavu. Nove riječi nastaju prema utvrđenom tvorbenom uzorku" [the word-formation pattern is the key term in the word-formation structure. New words are created on the basis of the established word-formation pattern]. (12) Barić argues further that the word-formation pattern is always binominal. In other words, on one side there is the base word, and on the other is the product of word-

formation product. Without this binary relationship the word-formation pattern does not exist. Thus, one can assume that the word-formation pattern exists if two conditions are fulfilled: on the one side there needs to exist the phonetic relationship between the base word and the word-formation product and on the other it is necessary to establish the meaning relationship between them. Thus, *čaša* ('glass') – *čašica* ('a small glass') is a word-formation pattern as it is possible to trace this binary relationship. Additionally, it follows that words from the same pattern have to fulfil the following conditions: the base word has to belong to the same morphological category and the word-formation device has to be the same. (Barić 1980: 12)

In the process of word-formation there is also another term that can be encountered, namely the word-formation type. Barić (1980: 13) argues that compared to the word-formation pattern, the term word-formation type is broader. What she means by this is the fact that the word-formation type includes words that are created in the same manner and with the same word-formation device. Thus the word-formation type includes words created in the same word-formation principle, while the word-formation pattern includes not only these words, but also those that are morphologically and semantically the same. For instance, compounds of the type *rukomet* ('handball') have more than one word-formation pattern which depends on the meaning of the compound.

In connection with the word-formation-pattern and the word-formation type, one should also mention the notion of analogy in the process of word-formation. According to Barić (1980: 13) there are two ways of creating new words. In the first way a new word is created on the basis of the transfer of the form and the content, i.e. *boravak* (n. 'stay') – *boravište* ('residence'), and in the second the word is formed by analogy, as in *raketodrom* ('rocketport') according to *aerodrom* ('airport'). Therefore, Barić (13) argues that it is necessary to introduce a new term, namely the analogical pattern. She argues that the process of word-formation is present in both patterns because it is not possible to talk about a word-formation product without it. The only difference is that the analogical and the word-formation pattern include different elements. The basic distinction is

u tome što se tvorbeni uzorak uspostavlja uz pomoć preoblike, a analoški je uvijek prisutan. Kad tvorimo novu riječ, mi, zappavo, tražimo

da li postoji već koja tvorenica sa istim odnosom dijelova, tražimo predstavnika takve tvorbe, tj. riječ koja će poslužiti kao ishodišna u procesu analogije, a to je upravo analoški uzorak [the fact that the word-formation pattern is established on the basis of the slight change, while the analogical pattern is always present. When a new word is created, we actually look for a word where we have the same relationship between the constituents, we look for the representative of that word-formation. In other words, it is the word that serves as the result word in the process of analogy, and exactly that is the analogical pattern]. (Barić 1980: 13)

When a new word is created, one should be able to interpret and understand it in an appropriate way. In order to be able to do this one should be able to differentiate between two levels of meaning: the lexical meaning and the word-formation meaning. Logically, the word-formation meaning only has those words that originate from the word-formation processes. The word-formation meaning is actually the altered meaning of the base word and it can best be described by the method of alteration. In this process the word-formation product is brought into the relationship with the word that was the starting point in the process of creating that word. (Barić 1980: 14) Additionally, Barić points out that all the words that have been created according to the same word-formation pattern have the same semantic or the meaning pattern. The semantic pattern consists of two parts: the first one is the word-formation product, i.e. the word that has been created in the process of word-formation, and the second is the description of that word. This description is gained from the process that she terms 'alteration'. This is the rule that is used to interpret a word-formation semantically. Barić argues that there are as many rules as there are word-formation patterns and that the general alteration is the one that can be used to interpret all the word-formations that have been created in the same way. This kind of alteration is the general semantic pattern and all the words that can be altered in the same or similar meaning and according to the one general pattern belong to the same semantic class. It can also be the case that a word-formation has more than one semantic patterns:

(33) jednospratnica ('one-story house'): *jednospratna zgrada* ('one-story house')
or *zgrada na jedan sprat* ('a house with one story')

In this case, the one meaning that sounds more natural to a native speaker or that is closer to the every-day way of expression is chosen, or the one meaning that

can be applied to more than one word-formations, and this is the usual way to analyse the given word-formation.

3.4.4 The role of motivation and half-motivation in the process of word-formation in Serbian

In word-formation it is necessary to distinguish between two terms: the base word (Serb. 'osnovna riječ'), that is the word that takes part in the process of word-formation and the word-formation product word (Serb. 'tvorenica'), that is the word that is created in the process of word-formation. (Barić 1980: 8) When it comes to the relationship between the base word and the product words, it is said that they are in the relationship of motivation. Thus she argues that

[o]snovna riječ u tvorbenom procesu motivira tvorenicu, tj. svojim izrazom i svojim sadržajem sudjeluje u ostvarivanju izraza i sadržaja tvorenice. Za riječ koja se može, izrazno i sadržajno, dovesti u vezu s drugom riječju kaže se da je motivirana, a za riječ koja se ne može, izrazno i sadržajno, dovesti u vezu s drugom riječju kaže se da je nemotivirana. [the base word motivates the product word in the process of word formation and with its form and content the base word takes part in the realization of the form and content of the product word. For a word that formally and contentwise can be brought into relationship with some other word, it is said that it is motivated, and for a word for which it is not possible to bring into any kind of relation with some other word it is said that it is unmotivated.] (Barić 1980: 8)

For instance, the relationship between *krov* ('roof') and *krović* ('small roof') can easily be recognized, as the derived noun *krović* can formally and contentwise be brought into relationship with the base word *krov*. When the relationship between the base word and the product word disappears, the word becomes unmotivated and thus is not a product of word-formation processes any more. For example *korica* in the meaning of *korica hljeba* does not have any diminutive meaning even though the suffix *-ica* denotes it. This is the point where the process of lexicalisation comes in. In other words, *korica* becomes lexicalised and thus cannot be analysed 'base (kor-) + suffix (-ica)', but rather as 'lexical morpheme (koric-) + grammatical morpheme -a¹⁹'. Apart from motivated and unmotivated words, there is a third type of words that Klajn (2002: 30-33) calls half-motivated words.

¹⁹ Nouns that end in -a in nominative singular belong to the 3rd category of the nouns in Serbian (Stanojčić 2004: 87)

As has already been discussed in the previous chapter, it can sometimes be the case that one part of the complex word is not transparent, in other words without meaning, or with that kind of meaning that becomes transparent only when one looks at the etymology of the word. This is not a phenomenon that is found in Serbian only. One of the most famous examples is probably *berry* in English. In words such as *cranberry* or *huckleberry* one can clearly distinguish between *berry* (the word that has a meaning) and the words *cran-* and *huckle-*. However, *cran-* or *huckle-* do not mean anything and cannot be found outside these complex words. In Serbian, it is the case with words that originate from the Turkish language. For example, in words such as *HALAbuka* ('extreme noise'), *PRAZIluk* ('leek'), *KREzub* ('toothless') it is clear that the second part is motivated, while the meaning of the first part sometimes cannot be found even with the help of an etymological dictionary. Additionally, examples can be found where the second part is unmotivated. Those are predominantly new loans from English or German. For example *šrafCIGER* (from German *Schraubenzieher*), *sesksEPIL* (from English *sex appeal*), *centarFOR* (from English *center forward*).

Klajn (2002: 32) also touches upon another solution. Namely, one could also say that words such as *seksepil* or *praziluk* are not complex words at all, because their second and the first part respectively are neither affixes nor full words. Thus, it may be assumed that *praziluk* is a simple word, because *prazi-* does not have a meaning in Serbian. However, if one tries to classify *praziluk* as a simple word it will not be possible to explain its relationship towards the motivated part of the compound, in this case *luk*²⁰. This relationship is not an accidental one, but it is based on the relationship of hyponymy. Further on, Klajn gives an example of *pakpapier* and *flispapir*. In *pakpapier* the first part of the compound is obviously motivated by the base *pak-* that one finds in the verb *pakovati* ('pack'). Thus, *pakpapier* is a compound word and it is a kind of paper. Therefore it is impossible to claim that *flispapir*, which is also a kind of paper is a simple word, just because of the fact that the base *flis-* does not exist in the Serbian language. Finally, Klajn (2002: 32) concludes that:

[n]eizbežan je stoga zaključak da složenice mogu biti i polovično motivisane, to jest sastojati se od jedne osnove sa leksičkim značenjem i druge koja nikakvog značenja nema, pa ipak u spoju sa onom prvom

²⁰ *Luk* means 'onion' in Serbian. Thus, *praziluk* is a kind of onion and therefore its hyponym.

daje određeno značenje složenici [it is inevitable to conclude that compounds can also be half-motivated. In other words, they can consist of one base with lexical meaning and the other one that does not have any meaning at all. However, in the combination with the first one it gives a certain meaning to the compound].

Thus, the role of half-motivation should not be disregarded and when it comes to the analysis of these words in this paper, they will be analysed as compounds and not as simple words. However, before we start to look at special types of noun compounds in Serbian, it will be necessary to clarify some terms that have been mentioned in this chapter and that sometimes can cause confusion in Serbian. The terms in question are root and base.

3.4.5 The notion of root and base

Root and base are terms that are mentioned in most of the grammars of Serbian. However, it is very often the case that they are mixed up and when it comes to the analysis of word-formation processes this differentiation might be crucial. When it comes to word-formation in Serbian, linguists usually distinguish between word-formation bases (roots and bases) and word-formation affixes. (Stanojčić 2004: 72) Affixes will not be discussed in greater detail, as their definition matches the definition of affixes in English to a greater extent. In contrast to affixes, word-formation bases require a short overview and the delimitation of the terms root and base. As already mentioned, word-formation bases are divided in two groups: roots (or root morphemes) and bases (or grammatical bases).

Root morphemes are those parts of words that carry the lexical meaning and cannot be split into smaller units that would also be carriers of the same meaning. For instance, in words *PIS-a-ti* ('write'), *PIS-a-li* ('we wrote'), *PIS-ac* ('writer'), *za-PIS* ('récord') the root is *-pis-*. In the group of words *ZNA-ti* ('know'), *ZNA-lac* ('expert'), *ZNA-k* ('sign'), *po-ZNA-t* ('famous') the root is *-zna-*, etc. As can be seen; roots are always monosyllabic and they can consist of only one vowel, as in *uz-E-ti* ('take'), of a vowel and a consonant, as in *ČU-ti* ('hear') and *PI-ti* ('drink'), of a vowel and two or more consonants, as in *VID-e-ti* ('see'), *ISK-a-ti*²¹ ('ask'), etc.

Grammatical bases "delovi su reči koji čuvaju vezu sa leksičkim (pojedinačnim) značenjem reči, tj. sa značenjem njenog korena, kao i sa gramatičkom kategorijom

²¹ Archaic.

(vrstom) dotične reči.” (Stanojčić 2004: 70) [are those parts of words that keep the relationship with the lexical meaning of the word, in other words with the meaning of its root, as well as with the grammatical category of that word]. In Serbian, there are two kinds of bases: the nominal (related to nominal words: nouns, pronouns, adjectives, numbers) and the verbal (related to verbs). When it comes to nominal bases, the most effective way to isolate them is to delete the ending for the genitive singular form, i.e. *sused-a* (‘from my neighbour’), *kuć-e* (‘from my house’), etc. Verbal bases are in so far complicated, as there are two types: the present and the infinitive base. The present base can be isolated when the ending for the 3rd form singular present is deleted, i.e. *peva-mo* (‘we sing’). In order to get the infinitive bases it is necessary to delete the infinitive ending *-ti*, i.e. *peva-ti* (‘to sing’). However, this distinction is only important for building tenses in Serbian and thus will not be a matter of further discussion in this paper.²² Additionally, it is important to mention that in some cases roots and bases have the same form, i.e. *kuća* (‘house’) has the same form *kuć-* for both the root and the base, while in other cases there is a clear distinction between what is the root and what is the base, i.e. *pleme* (‘tribe’) the root is *-plem-*, while the grammatical base is *plemen-*.

According to Stanojčić (2004: 70) the reason why one should always bear in mind this distinction between roots and bases is the fact that compounds in Serbian can be composed of two or more word-formation bases²³ or two or more words from the same or different grammatical category, and these words can be either simple or complex, i.e. *Beograd* (< *beo* + *grad*), *parobrod* (‘steamboat’) (< *para* + *brod*), etc. In building these compounds, it is necessary sometimes to insert certain vowels between word-formation bases (as in *par-o-brod*) for reasons of easier pronunciation. This vowel is called the ‘binding vowel’ (Stanojčić 2004: 74; Klajn 2002: 23) or the ‘binding format’ or just ‘binder’ (Babić 1986: 30).

3.4.6 Shortening of bases and language economy

In the process of word-formation the notion of language economy should not be forgotten. Namely, products of word-formation should be optimal when it comes to the number of syllables and the amount of meaning that the word has. For instance, just as in English, compounds in Serbian that have more than two bases

²² For further discussion about verbal bases see Stanojčić (2002: 67-74)

²³ Words-Formation bases are roots and grammatical bases

can also be reduced to binary constructions. Let us have a look at the following example: *staro visokonjemački* ('Old High German'). The word is derived from *stari* ('old') + *visoki* ('high') + *njemački* ('German'). Thus if we analyse it as a hierarchical structure that involves binary elements, the above mentioned word can be represented using the bracketing system in the following way:

(34) [staro[visoko] njemački]

As it can be seen, the three-base word is reduced to a binary construction in order to achieve the optimal cohesion of the compound. However, this reduction to a binary construction very often results in the shortening of the bases. In most cases the original base is taken from a suffixal derivative, although the presence of suffix is necessary for the meaning. For example, the word *toplota* < *toplo-ta* ('heat') is shortened in the word *toplomjer* ('thermometer') and thus the form **toplotomjer* is not attested. The same happens in *visoravan* ('plateau') < *visok* ('high') + *ravan* ('plane'). Suffixes are sometimes deleted in words where there is not any saving of syllables and this happens only for the purpose of easier pronunciation of words, i.e. *rimokatolik* ('Roman Catholic') < *rimski* + *katolik* instead of **rimskokatolik*, *grkokatolik* ('Greek Catholic') < *grčki* + *katolik* instead of **grčkokatolik*, etc. There are also a lot of compounds with shortened bases that have foreign origins, i.e. Russian as is *socrealizam* ('sociorealism') *socijalni* + *realizam*, German as is *profi-fudbal* ('professional football') < *profesionalni* + *fudbal*, etc.

3.4.7 The role of the binding vowel

The binding vowel (also called 'the binding morpheme') is an issue that many linguists have dealt with (Barić 1980; Belić 1986; Klajn 2002). The basic question that all of them tried to answer is whether this vowel is a kind of interfix, or a morpheme just like a prefix or suffix, or a morphonological element with pure formal function of binding. Rammelmeyer (1975) uses the term 'Kompositionsmorphem'. However, he states that in contrast to prefixes and suffixes "besitzt [...] keinen eigenen semantischen Wert." This idea is also supported by Baretić et al (1979: 225): "[s]pojnik nije značljivo tvorbeno sredstvo [...] ali i spojnik je značljiva jedinica – on znači spajanje, samo što nije tvorbeni nego gramatički sadržaj" [the binding vowel is not a word-formation unit with

meaning [...] however, it is a meaningful unit. The important fact is that it does not have any word-formation, but grammatical content]. As Klajn (2002: 24) argues this formulation cannot help one to come to any further conclusions because when it comes to forming compounds it is not possible to talk about content that is grammatical, but not word-formational.

The most important thing is to decide how many binding vowels there are in Serbian. Stevanović (1964) states that there are two: *o* and *e*. However, this topic is hardly mentioned in the chapter about compounds. Nevertheless, he argues that the vowel *o* can be found before palatal consonants, as in *bilj-o-jed* (< *biljka* ('plant') + *jesti* ('eat')). Additionally, he mentions that formations such as *srednj-e-školski* ('middle school'), *prednj-e-nepčani* ('palatal') are very common even among scholars. However, in dictionaries and grammar it is stated that formations *srednjo-* and *prednjo-* should be preferred. When it comes to the binding vowel *e*, Stevanović states that there are only few compounds with only this vowel and exemplifies it with the noun *oceubica* ('one who kills his/her father'). Again, Klajn (2002: 24) says that the number of compounds with *e* is much higher than Stevanović assumed and gives the following examples: *očevidac* ('witness'), *moreplovac* ('sailor'), *kraljeubica* ('the one who kills kings'), etc.

Belić (1986: 30) states that the status of the binding vowels *o* and *e* is not a matter of question. Just as Stevanović (1964: 116-118), he argues that *o* can be found before palatal consonants and he also points out that there are cases where some people use *e* and others *o*, even though according to grammars and dictionaries *o* should be preferred (*srednj-o-školac* and *srednj-e-školac*). What is innovative about Babić's theory is that he introduces *i* and *u* as binding vowels. He states that *i* comes if the first part of the compound is a verb, as in *cjep-i-dlaka*, *rasp-i-kuća*, *vuc-i-batina*, etc. He supports this analysis by stating that the first part of the compound in these cases is either the present or the infinitive base. However, Klajn (2002: 24) agrees with Rammelmeyer (1975: 93) who argues that all the compounds of the type V + N that at first sight have *i* as a binding vowel, can actually be interpreted in another way. This *i* is not a binding vowel, but it is the form of the 2nd person singular imperative. According to Klajn (2002: 24) the only example that supports the proposal of *i* as a binding vowel is *cjepidlaka* ('hairsplitter') (< *cijepati* ('tear') + *dlaka* ('hair')), since the imperative form of *cijepati*

is *cijepa-j*. Otherwise, all the other examples clearly exemplify the imperative *i*. For instance, *raspikuća* ('spendthrift') is derived from *raspi* ('spend') and *kuca* ('house'). The binding vowel *u* is "ograničen [...] samo na neke riječi u prvom dijelu" [restricted only to some words in the first part] (Babić 1986: 30), as *bratu-* ('brother'), *polu-* ('half'), *hiljadu-* ('thousand'): *bratućed*, *polusprat* ('half-floor'), *hiljadugodišnji* ('millennial'), etc. It has to be stated that Barić (1980: 18-19) does not see *polu-* as adjective, but rather as prefix and thus all the words created with it are a part of prefixation, and not compounding. Additionally, Klajn (2002: 42) claims that *bratu-* is the dative form of *brat*, which brings us to the conclusion that there are very few examples where one can find *u* as the binding vowel. Thus he concludes that *i* and *u* should not be counted as binding vowels, as very few examples can be found. Finally, there are authors that write about the zero binding vowel. For instance, Rammelmeyer (1975: 93-94) points out that the zero vowel is the most common in the German-Turkish noun-noun compounds, as in: *duvankesa* ('tobacco pouch'), *paradajz-čorba* ('tomato-soup'), etc. However, he also states that this opinion can be denied "[f]alls man jedoch nicht bereit ist, ein Kompositionsmorphem -Ø- anzunehmen, muß man von einer Zusammenrückung von Stamm + Substantiv [...] ausgehen." (93) Babić (1986: 30) does not mention the zero binding vowel, but he writes about compounds without binding vowel. Barić (1980: 24) also writes about the binding vowel (Serb. *pojnik*) which can either be present, as in *par-o-brod* ('steam boat'), or hidden, thus a zero binding vowel, as in *duvan-Ø-kesa* ('tobacco pouch'). Barić also mentions that some Russian linguists proposed the term *interfix* for the binding vowel, because the Latin word *inter* means 'between' and thus the term *interfix* expresses the fundamental role of the binding vowel. When it comes to the role of the binding vowel, there are two major opinions among Russian linguists about what it actually is. One opinion supports the idea that the binding vowel is just an inserted vowel without any meaning, that received the name *interfix* analogically to the terms prefix and suffix. However, as the prefixes and suffixes have meanings, it is not possible to claim that the binding vowel is a kind of affix. The other opinion goes a bit further and linguists that support this idea claim that the binding vowel should be named binding morpheme. However, as a morpheme is the smallest meaningful unit in a language, one could also assume that a binding vowel also has certain meaning and this is not the case. Thus, when it comes to compounds

where the binding vowel is not present, it will be assumed that these compounds are without binding vowel, and not the binding vowel -Ø.

To conclude the discussion about binding vowels in Serbian, it must be said that the only vowels that can be claimed to be the binding ones are *o* and *e*. Furthermore, it must be stated that *o* is by far more common, not only in Serbian, but also in other European languages. According to Klajn (2004: 26) the greatest part of Greek compounds had the first part that ended in *o*, while in other European languages -*o* serves as the binding vowel in some adjectives: Engl. *spatio-temporal*, Fr. *latino-américain*, It. *latino-americano*, etc. And finally, as it has been shown, the claim that something like a zero binding vowel exists, is not justified and thus will not be dealt with as a binding vowel in this paper.

4. Types of compounds in English

When it comes to the distinction of the types of compounds in English, it gets really difficult to find one straight division for all of them. As we will see later, it is not even possible to say whether certain words are compounds or not. The main problem here is how one analyses the compounds. Different analyses (morphological or semantic-syntactic) will provide different types of compounds. For this reason, it is necessary to have a look at some of the authors that have dealt with this topic. As it has already been stated at the beginning of this thesis, Hans Marchand is seen as the father of European morphology, and therefore it is suitable to start with his remarks about compounds. Before doing this, it is important to clarify some of the terms that he introduced.

4.1 Endocentric compounds

The thing that all morphologists agree on (Marchand 1969: 11, Plag: 2003: 145, Lieber 2005: 375) is that compounds that have their heads within the compounds and the head is obligatory the right-hand member (determinatum) of the compound are called endocentric compounds. The element *endo* means 'inside' and *centric* comes from 'centre'. The most important feature of these compounds is that they "denote a subclass of the referents of the head." (Perkles 2008: 23) Expressed in other terms, endocentric compounds are nothing but hyponyms of their heads.²⁴ This is sometimes called *Is A Condition* (Allen 1978) or *type-of test* (Bauer & Huddleston 2002: 146) Hyponymy seems to be the easiest and the most helpful way to describe the N + H compounds in the modifier/modified way, however, "neither hyponymy [...] nor straightforward interpretation [...] are inherent features of compounding". (Fernández-Domínguez 2008: 72)

4.1.1 Expansions

The principle of combining two words arises from the natural human tendency to see a thing identical with another one already existing and at the same time different from it. (Marchand 1969: 11)

As Marchand states, it is a fact that humans have a natural need to create new words in order to be able to describe the world around themselves. For that reason, the fact that some word-formation processes are extremely productive

²⁴ For further discussion see Kastovsky 2008

should not be surprising at all. When it comes to compounds, some linguists (Plag 2003: 132) go so far as to claim that this is the most productive way of creating new words in English. Let us analyse the following examples: *coffee machine*, *colour-blind*, *refill*. If we take the word *coffee machine*, it can be stated that the identity is expressed by the word *machine*, and the difference by the word *coffee*. Now, *coffee machine* compared with *machine* is a modified, expanded version of *machine*. Further on, the word *coffee machine* will be found in basically the same contexts as the word *machine* and it will also retain the same syntactic features; they both belong to the word class that we know as nouns. The next example, *colour-blind*, can be interpreted in the same way; it is the expansion of the adjective *blind*. A person is called *colour-blind* because he or she is seen as blind, but only with regard to colours. The verb *refill* is basically the verb *fill* and in most of the cases they are exchangeable except for the modification expressed by *re-*. However, this does not affect the class of the word, which is that of a verb. This type of example is not going to be dealt with in this paper, since it belongs to another class of word-formation, namely derivation.

Marchand (1969: 11) argues that combinations of the types *coffee machine*, *colour-blind*, and *refill* are morphological extensions of the words *machine*, *blind*, and *fill* respectively, and therefore they are termed as **expansions**. “An expansion will then be defined as a combination AB in which B is a free morpheme (word) and which is analysable on the basis of the formula $AB=B$.” (11) This means that the combination AB belongs to the same word class to which B belongs. Combinations AB that contain free morphemes for both determinant and determinatum (*coffee machine*, *colour-blind*) will be termed compounds. Combinations AB, in which the determinatum is a free morpheme, while the determinant is a bound morpheme, are prefixed words and therefore not a part of the analysis in this paper. Thus, both compounds and prefixed words are subgroups of the larger class called ‘expansions’.

4.1.2 Synthetic compounds

Synthetic compounds, sometimes also called verbal (Lieber), deverbal or verbal nexus compounds, are combinations whose second elements are deverbal derivatives from verbs which form a direct syntagma with the determinant (e.g.

car-driv-er, *watch-mak-er*). According to Marchand (1969: 16), they are nothing but derivations from a verbal nexus. However, since they are analysable by the morphological pattern N + N, they can be treated as compounds (formally speaking).

- | | |
|---------------|---------------------------|
| (35) N+(V-er) | a. watchmaker, bookseller |
| | b. nut cracker |
| | c. party drinker |
| N+(V-ing) | d. house keeping |
| | e. heart-breaking |

When it comes to the structure of the examples in (35), it can be argued that these compounds are endocentric compounds since their head is to be found in the compounds. They can also be analysed according to the formula AB=B, which could make them look like expansions. At first sight one can come to false conclusions, put them in the same group with the primary compounds (expansions), however, it can be observed that they do not actually fit all the rules that an expansion should fulfil. Namely, as it has been stated earlier, expansion are created of two independent lexical items, and the head is obligatory independent. Obviously, this is not the case in (35). The first part of the compound is an independent lexeme, the second part is not. In *housekeeping* the first element is a common noun, while the second element *keeping* is not that kind of a lexical unit (it usually does not exist as a lexical word). In the same manner, *maker* in *watchmaker* does not exist as an independent word. Marchand termed them as 'functional derivatives', meaning that they only represent the syntactic relation of the underlying paraphrase. (1969: 16) Analogically, they should not appear in a dictionary since they are not full semantic units. Additionally, from the morphological point of view these composites are compounds, but from the syntactical standpoint, they are nothing but derivatives.

- (36) a. morphological analysis: watch/maker (N + N), church going (N + N)
 b. syntactic analysis: watchmak/er (O-P-S), churchgo/ing (O-P-Pn)

It can be seen from (36) that these examples actually satisfy conditions for a compound, which means that they are made up of two lexemes. However, they look like suffixed formations on the level of the underlying paraphrase.

As already stated, there are two types of syntactic compounds: the *watchmaker* and the *housekeeping* type. (Marchand 1969: 16) The type watchmaker is found to

be a derivative from the verbal nexus 'one who makes watches'. In order to make a combination *watchmaker*, the word order of the verbal nexus has been inverted to achieve the N + N pattern, and as already said, it makes what "is a derivation at the level of the underlying sentence into a syntactic compound at the level of morphological structure." (16) The compounds of this type refer to someone whose role is that of agent, meaning that someone does whatever the verb signifies. (Perkles 2008: 29) Marchand further divides this group into two subgroups, namely, the type *deer hunter* and the type *watchmaker*. In the type *deer hunter*, we actually deal with two independent lexical items, *deer* and *hunter*, which form N + N compound and the compound is analysable as 'hunter of deer'. The question is now, how can we account for the fact that they are on the one side composed of two independent lexical items, but on the other seen as synthetic compounds? Marchand proposes the following solution: the following combinations belong to this type: *ballet dancer*, *cigar smoker*, *crime reporter*. Even though their second element is an independent lexical item, this does not prevent them from being analysed as verbal nexus combinations. A *cigar smoker* is 'a smoker of cigars' and 'one who smokes cigars', a *ballet dancer* is 'a dancer of ballet' and 'one who dances ballet', etc. This type is one of the most productive types of Modern English word-formation. (Marchand 1969: 80)

Until now, synthetic compounds have been treated as compounds because of their morphological structure. However, Marchand claims that there is a great difference between primary and synthetic compounds. As it has been pointed out, synthetic compounds are compounds only on the morphological level. In contrast to primary compounds, which consist of purely nominal constituents (e.g. *armchair*, *stone wall*), whose immediate constituents are independent elements at the level of the underlying sentence ('wall made of stone' is stone wall), in syntactic compounds the second morphological element (*maker* in *watch maker*) is an S-P group at the level of the underlying sentence (he makes) "whose nominalized form is represented by *maker* at the level of morphologic structure." (Marchand 1969: 18) What is important here is the fact that at the level of underlying sentence synthetic compounds are explicit, while the primary compounds are implicit syntagmas. This is relevant for the explanation of syntagmas. While all compounds can be explained from the syntactic relations underlying them in sentences, synthetic

compounds are reduced from full sentences, and in that way represent the entire “verbal nexus in an overt form”, (18) with the predicate being the essential part of the sentence, thus allowing us to interpret the relation between the constituents of the compounds in an accurate way. However, the problem arises when there is a compound (e.g. *chain smoker*, *day dreamer*, or *party drinker*) where the modifying element (determinant) cannot be seen as the object of the underlying sentence. A *day dreamer* is not someone who dreams days, a *party drinker* is not a person who drinks parties, and a *chain smoker* is not someone who smokes chains. These nouns would rather have the function of adverbials for time, place, and manner respectively. Thus their deep structure would be as follows: day dreamer AdT-P-S, party drinker AdP-P-S, chain smoker AdM-P-S. As it can be seen, even though *party drinker* and *day dreamer* belong to the same group as *watch maker* or *letter writer*, it is not possible to analyse them in the same way. Oshita (1994: 188) even claims that the *day dreamer* group should not be treated as synthetic compounds at all. He claims that

the only general observation we can make is that [...] such a noun quite freely occurs on its own or as a head of a compound and that, when it is a compound head, its nonhead constituent may or may not appear to superficially satisfy the argument structure requirement of the base verb. (Oshita 1994: 188)

In other words, this would mean that these instances are pure accidental syntagmas, and depending on the type of nouns that form the syntagma, it would be possible (or not) to establish the relationship between the constituents. Oshita mentions the so called argument structure of compounds and therefore at this point it is necessary to introduce this approach to the analysis of compounds.

In order to make the interpretation of synthetic compounds complete, and explain the possible meaning relationships between compound constituents, it is necessary to mention the theory of the so called missing arguments (Carstairs-McCarthy 2002: 63, Lieber 1983, Plag 2003: 148). In order to make this clear, let us compare the following examples: *hair restorer* and *apple eater*. The second element of the compound is a noun derived from a verb and thus the whole syntagma fits the class of synthetic compounds. As Carstairs-McCarthy (2002: 62) claims “[v]erbs, unlike most nouns and adjectives, impose expectations and requirements on the noun phrases that accompany them in the sentence.” In other

words, with intransitive verbs, such as *sleep*, we expect to find one noun phrase as subject, with *restore* and *eat* we also expect to find a noun phrase as object, and with the verbs *give* or *receive* we find a third indirect object noun phrase. These expected noun phrases are called ‘arguments’. Let us now come back to the example *hair restorer*. Relying on the missing argument interpretation, the most natural way to interpret the compound is quite precise: the first element expresses the object argument of the verb. To generalize, an X-restorer, whatever X is, is something or someone that restores X. In our example, X is hair. This whole process by which a phrase or a word in the neighbourhood of a head word is given the status of the head word’s argument, is called ‘argument linking’. (Plag 2003: 149)

However, argument linking in synthetic compounds sometimes fails. This happens if the first constituent of the compound is not semantically compatible with its role as argument. For instance, *Sunday driver* is not a person who drives a Sunday, but someone who drives on Sunday. In the same manner, a *street seller* is not someone who sells streets, but someone who sells on the street. How could we account for these irregularities? One possibility would be to say that in order to arrive at the precise meaning of compounds, it is not only the pure linguistic knowledge that is necessary, but one should also take our world knowledge into account. Thus, relying on our world knowledge, we would not interpret Sunday driver as someone who drives a Sunday, but rather as someone who drives on Sunday, because our knowledge of the world blocks the first interpretation. Therefore, it can be concluded that the interpretation of synthetic compounds (but also of primary compounds as we will see later²⁵) is extremely variable and that it depends on the argument structure of the head, the semantics of the two nouns, the possible conceptual relationships between the two nouns, as well as on the surrounding discourse.

4.1.3 Primary or root compounds

As has been pointed out by Marchand (1969: 80), Jespersen (1909-1949, vol. IV: 143-145), Bauer (2003: 202-207), the type N+N is the most productive way of creating new compounds. Thus, it is to be expected that there are many different

²⁵ See chapter 4.1.3

ways of joining two nouns together and thus this chapter will deal with the analysis of compounds that are composed of two nouns, sometimes called in literature 'primary' (Marchand: 1696: 11) or 'root' (Lieber 2005: 375) compounds.

Primary compounds (in contrast to synthetic compounds that will be discussed in the next chapter) are "expansions of which both the determinatum and the determinant are words". (Marchand 1974e: 323-324) It has also been stated that they are analysable by the formula $AB=B$. Compounds such as *steamboat*, *stonewall*, *moonlight*, or *armchair* can be interpreted in this way and therefore can be called primary compounds. In all these cases "the determinatum represents the element whose range of applicability is limited through the determinant". (Marchand 1974e: 324) Let us have a closer look at their semantic analysis. *Steamboat* is a kind of a boat that is powered by steam, *stonewall* is a kind of a wall that is made of stones, *moonlight* is a kind of light that is radiated from the moon, and finally *armchair* is a chair in which you can rest your arms. However, this is a very broad and general description. Without any context, how would it be possible for the speaker to know that *moonlight* is the light that is radiated from the Moon, and not a kind of light that has the shape of the moon? The same applies to the rest of the compounds. Olsen (2000a: 910) claims that all compounds are potentially ambiguous but "the ambiguity is meaning-driven and contained within the limits of plausible connections between the constituent meanings." This would mean that that a person can derive the intended meaning of the compound from the context in which the compound has been mentioned. Kay & Zimmer (1990: 239) argue that the way one interprets a nominal compound may be "extended to limits of one's ingenuity", since they do not contain a verb, as it is the case with synthetic compounds, and thus they are always potentially ambiguous. (Kastovsky 1986: 72)

As has been stated earlier, the formal head of the compound serves as its semantic head as well, while the first constituent has the function of its modifier or complement. Some linguists (Marchand 1969: 18) claim that all compounds can be explained "on the basis of syntactic relations underlying them in sentences". What Marchand means by this statement is that compounds are created on the basis of underlying sentences and the use of a topicalization rule. This rule picks up the topic and assigns it the status of the head (determinatum) of the compound.

Further reduction of the sentence isolates the non-head (determinant) and allows it to expand the determinatum into a compound syntagma. (Olsen 2000a: 906). In this way both members of the compound retain their grammatical functions. What is interesting about primary compounds is the fact that a verbal nexus is not overtly expressed, but it is implicit. Let us have a look at the following examples:

- (37) a. car thief = 'someone who (S) steals cars (O)'
 b. paper basket = 'basket into which (AdP) one throws paper (O)'
 c. cave man = 'man who (S) lives in a cave (AdP)'
 d. wall paper = 'paper (O) which is put on the wall (AdP)'

As can be seen, the verbs in example (37) are implicitly present in the nouns and the relation between the constituents of the compounds is assumed. Usually, these verbs are generic ones, which means that there is a certain, limited number of these verbs (e.g. *live*, *give*, *put*, etc.).²⁶ From this point of view it can be stated that primary compounds can have many types of references, and that their determinata can have the function of subjects, objects, adverbials, etc. The fact that one interprets *cave man* as 'man who lives in a cave', and not as 'man made of cave' or 'man who is a cave', has to do with the fact that we know how the world works, and thus interpretations such as 'man made of a cave' or 'man who is a cave' are blocked.

As has been stated earlier, a compound is interpreted "by relating the two members of the compound to each other in terms of the typical relationship between the entities referred to by the two nouns." (Plag 2003: 148) When it comes to this 'typical relationship', Plag argues further that it is necessary to distinguish between two different classes of nouns, 'sortal' and 'relational' nouns. Sortal nouns are used for classifying entities. For example, entities might be called *chair*, *desk*, *table*, etc. Relational nouns denote relations between a specific entity and a second one. Thus, someone cannot be a *daughter* without being the daughter of someone. In the same manner, one cannot perform a *surgery* without performing it on something. This second, conceptually necessary entity, to which a relational noun relates, is called 'argument'.²⁷ To come back to the problem of the interpretation of primary compounds, it can now be stated that if the right-hand member of the compound is a relational noun, the left-hand member will normally

²⁶ For further discussion see Lees (1970)

²⁷ See chapter 4.4 for further discussion about argument

be interpreted as the argument of the relational noun. Thus, *brain surgery* would be analysed as 'surgery performed on the brain', *finger surgery* as 'surgery performed on fingers, etc. Again, as it has already been the case with the synthetic compounds, argument linking can sometimes fail. This happens if the two members of the compound are semantically incompatible. For instance, *computer surgery* is not interpreted as surgery performed on computers, because computers are not usually treated by surgeons in the way human organs are. If this interpretation is impossible, a new one should arise, one that relies on the possible relationship between the constituents. According to Plag (2003: 150), the inferencing procedure is likely to happen. As we know, computers are used in all kinds of medical instruments, and some medical instruments are used by surgeons, so another interpretation becomes possible, 'surgery performed with the help of a computer'.

Another problem can arise whenever a non-relational noun becomes the head of the compound. For instance, *stone wall* is likely to be interpreted as 'wall made of stones', because it is a typical relationship that walls are made of stone. However, depending on the context, this interpretation is not the obligatory one. One could interpret it quite differently, for example as a wall against which a stone was flung, or a wall that is painted with a graffiti showing a stone, etc. A similar example is the compound *marble museum*. There are two possible interpretation of it, again depending on the context. The first interpretation has to do with the concept of the museum as a building. A building is made of stone, marble is a kind of stone used for constructing buildings, and thus *marble museum* can be interpreted as 'museum built of marble'. Another interpretation is based on the concept of a museum as a place where precious objects are kept. As marble is a kind of expensive stone used for making sculptures, for instance, a *marble museum* could be a museum where marble objects are exhibited. From these examples, it can be seen that the interpretation of nominal compounds depends on "possible conceptual and semantic properties of the nouns involved and how these properties can be related to create compositional meaning in compounds." (Plag 2003: 150)

The *marble museum* example brings us to another major factor that is involved in the interpretation of nominal compounds, namely the surrounding discourse. The

way one interprets *marble museum* largely depends on the preceding discourse. For instance, if the word occurs in a newspaper article about an exhibition of marble sculptures, it is very probable that we will interpret it as a place where marble objects are displayed. In contrast, in a context about building materials of public buildings, the interpretation 'museum made of marble' is more likely. (Plag 2003: 150-151) There are also examples where certain pragmatic factors lead to the lexicalisation of a word-formation, so that only one meaning is usually accepted. (Perkles 2008: 25) Let us have a look at the following examples:

- (38) a. snake poison = 'poison produced by snakes'
 b. rat poison = 'poison for killing rats'

By analogy with *rat poison*, one could interpret *snake poison* as 'poison for killing snakes'. However, this meaning is blocked due to pragmatic factors, which is that many snakes are venomous, even though such an interpretation is not excluded. Bauer (1979: 46) argues that snake poison

Cannot be accounted for solely in terms of the lexemes involved, since not all snakes are venomous. Thus it is only because of the pragmatic knowledge that some snakes are venomous that the speaker/hearer can assign the interpretation 'venom from a snake's glands' to *snake poison* as opposed to the interpretation 'poison for killing snakes' [...]

In connection with this, it is necessary to mention Coseriu's proposal. He (1977: 51) argues that for the analysis of a compound, for instance *paper basket*, the only thing that is known is that the relation between the constituents is of a very general nature, i.e. it is just "eine allgemeine ‚präpositionelle‘ Funktion". Accordingly, at the level of the linguistic system the compound *paper basket* has only the meaning 'basket which has something to do with paper'. The reason why the unambiguous interpretation such as 'basket for paper' or 'basket made of paper' arises is due to "die allgemeine und [...] die spezifische – mit Kontext und Situation zusammenhängende – Sachkenntnis". (Coseriu 1977: 51) The fact that *paper basket* is usually interpreted as 'basket for paper' is attributed "to a referential fixation at the level of the norm of the language, i.e., it has in fact been lexicalised in this function." (Kastovsky 1986: 71) At this point it can be stated that lexicalisation has a kind of a disambiguating function. (Kastovsky 73) Due to certain extralinguistic factors one of the numerous possible meanings of a word-formation is lexicalised for the whole speech community, and thus blocking

another possible meaning, but not excluding the theme completely. According to this, *snake poison* is usually interpreted as ‘poison produced by snakes’, while *rat poison* is usually interpreted as ‘poison for killing rats’. However, in certain situations, other interpretations, such as ‘poison for killing snakes’ would also be possible. Kastovsky (73) concludes by saying that they are not vague but “systematically ambiguous”, and that pragmatic factors, such as lexicalisation, for instance, do not create interpretations, but they rather “disambiguate compounds by selecting one potential reading as the one fitting the given context.”

In order to accomplish a complete analysis of primary compounds, it is necessary to have a look at one additional factor – stress. As has already been stated, compounds usually have forestress, which means that the left-hand constituent has the strong stress, while the right-hand element has the weak one (e.g. *college student*, *dinner hour*, *flu season*, *instruction book*, etc.). According to Olsen (2000b: 60)

[l]efthand stress [...] occurs in primary compounds where a relational notion is not overtly expressed by the head noun, but is inferable on the basis of the meaning of one of the two constituents. The inferred relation is then used to combine the meanings of the two elements in the compound.

In order to make this clear, let us have a look at the following examples: *stock market*, *space scientist*. The interpretation of the compound *stock market* is based on the relation ‘deals in’ that is derived from our knowledge of what ‘markets’ usually do. Thus, we interpret *stock market* as ‘deals in stocks’. Similarly, our knowledge tells us that scientists are usually engaged in studying something, hence we would interpret *space scientist* as ‘scientist studying space’. This interpretation holds for most of the N+N compounds. However, Olsen points out that there are examples where right-hand stress is found. She claims that this is typical of certain semantic groups. It is usually the case when the first constituent functions as a temporal or locative modifier (e.g. *hotel kitchen*, *ocean habitat*, *world travel*, *spring showers*, *summer dress*, *Sunday drive*, etc.); when the first constituent denotes the material out of which something is made (e.g. *cotton blouse*, *glass window*, *wool coat*, *silk tie*, *copper wire*, etc.); when the relation between the constituents is a predicative one (e.g. *surprise encounter*, *giant telescope*, *senior scientist*, *woman doctor*, *key word*, etc.).

4.2 Copula compounds

The fourth major group of compounds in English that will be discussed in this paper is the so called copula compounds. In contrast to the endocentric compounds that can be analysed on the basis of the formula $AB=B$, and the exocentric compounds that have the structure $AB\neq B$ and $AB\neq A$, the copula compounds can be interpreted by the mathematical formula $AB=A$ and $AB=B$. This means that the compound as a whole is equal to its determinant, as well as to its determinatum. The name of this group comes from the fact that the verb *to be* appears twice in the underlying sentence that is used to paraphrase the compound. For example, the compound *girl friend* can be analysed as ‘someone who is a girl’, as well as ‘someone who is a friend’. In Sanskrit grammarian terms of this type of compounds are called ‘dvandva’. (Plag 2003: 146-147) In order to be able to analyse morphological and syntactic semantic features of copula compounds, let us have a look at the following examples (Olsen 2003c: 295-300):

- (39) a. oak tree, murder robbery, washer dryer, camper trailer
b. banking businessman, managing director, broker analyst,
c. filmmaker-playwright, movie star-singer, manufacturer shipper
d. scholar-deputy, professor-consultant
e. reporter-narrator, author-philosopher,
f. hacker-programmer, user-programmer
g. developer-architect, designer-builder
h. doctor-patient gap, brother-sister relationship, father-sun duet, teacher-parent council

On the basis of the examples listed above, it can be stated that most of the copula compounds have the morphological pattern N+N. When it comes to the constituents, they can be either simple, as in *oak tree*, or there are cases where both of the constituents are complex lexemes, as in *filmmaker-playwright*. Additionally, Olsen (2000c: 295) argues that stems of the copulatives can contain derivational suffixes (such as *-er* in *reporter* and *narrator*), but usually do not have stems that are derived by prefixation. (e.g. Kennerly was *co-writer* and *co-executive producer* vs. **co-writer-co-executive producer* Kennerly).

When it comes to the morphosyntactic properties of copula compounds, Olsen (2000c: 292) argues that they can be represented as endocentric compounds, as they take their morphosyntactic properties on the rightmost constituents. In these terms, this could be called the head of the compound.

- (40) a. filmmaker-playwrights, designer-builders
 b. doctor-patient-gaps, teacher-parent councils (Olsen 2003c: 295-300)

Further on, Olsen states that this is not the only morphosyntactic property that makes these compounds similar to endocentric compounds. Namely, copula compounds show the recursive (binary branching) structure, a feature that is typical for endocentric (primary) compounds. Let us have a look at the following examples (Olsen 2003c: 295-300):

- (41) a. actor-writer-impressionist
 [[actor writer] impressionist]
 b. artist-writer-film creator
 [[[artist writer] film] creator]
 c. dancer-choreographer-actor-designer
 [[[dancer choreographer] actor] designer]

As it can be seen from the examples, all the compounds follow the hierarchical pattern that are typical of endocentric compounds. Olsen also points out a further lack of difference between the endocentric and copula compounds. She states that it can be found in the “complex recursive interaction between the determinative and copulative subconstituents in the complex structures.” (293-294)²⁸ Let us analyse the following examples (Olsen 2003c: 295-300):

- (42) a. brain size-body size proportions
 [[[brain size][body size]] proportions]
 b. brain-body size trend
 [[[brain body] size] trend]
 c. brain-language-vocal tract coevolution
 [[[brain language] [vocal tract]] coevolution]
 d. comedy writer-director-producer
 [[[comedy writer] director] producer]
 e. eye-mind-mouth connection
 [[[eye mind] mouth] connection]

In a. two endocentric compounds (*brain size* and *body size*) are combined to form a complex copulative compound (*brain size-body size*) and then it is embedded in an endocentric relation with the noun proportions. In b. the copulative compound brain-body is embedded in an endocentric relation with size, and then this whole complex unit is again embedded in an endocentric relation with trend. In c. the endocentric compound vocal tract forms “a hierarchically structured three stem copulative compound with brain and language, then this complex copulative is

²⁸ Olsen uses the term ‘determinative compounds’ for endocentric compounds

embedded in a determinative relation with coevolution.” (294) The same analysis could be applied to the examples in d. and e. Finally, Olsen concludes that when it comes to the structure of copulative compounds, it seems that the stems combined to form a copulative compound can be either copulative or determinative complexes, as long as they themselves consist of two stems. However, Olsen (2003c: 295-300) also states that tripartite stems, such as newspaper baron, film studio mogul and television-network owner are too complex to serve as constituents in a copulative compound and therefore the explicit syntactic coordination is preferred.

- (43) a. (Murdoch,) a newspaper baron, a film mogul and a television-network owner
 b. *film-studio mogul-television-network owner (Murdoch)

It has been stated above in the chapter that stems derived by prefixation are not usually put together to form a copulative compound. Instead they also have the structure of an explicit coordination. (e.g. Kennerly was *co-writer* and *co-executive producer* vs. **co-writer-co-executive producer* Kennerly).

When it comes to the semantic properties of the compounds, it has to be stated that the semantic features of copulative compounds depend on the type of relations between the constituents. According to this relation, copula compounds are divided into two groups. The first group is represented in (39) a.-g. In all these examples “[e]ach form [...] refers to one entity that is characterized by both members of the compound.” (Plag 2003: 146) For instance, a *reporter-narrator* is someone who is at the same time a reporter and a narrator; an *author-philosopher* is an author who is also a philosopher. It seems that in these cases neither of the constituents is semantically prominent, and it can be argued that they both equally contribute to the meaning of the compound. Since it is claimed that the relationship between the constituents is one of apposition, these compounds are called ‘appositional compounds’. (Plag 2003: 146) According to Marchand (1969: 40-41) this group can be subdivided into two smaller groups. *Oak tree* and *teaching profession* are the representatives of the so called ‘subsubstantive compounds’. What he means by this term is that in these combinations A is or is assumed to be a logical subclass of B. Thus, in *oak tree*, *oak* is by definition a *tree*, with *tree* representing the genus and *oak* species. On the other hand, in *teaching*

profession, *profession* is not the logical genus of *teaching*, “but is only ad hoc assumed to do so as one possible class we chose to consider for our purpose.” (Marchand 1969: 41) *Teaching* can also be a pastime, a vocation, or some other thing, while an *oak* can only be a tree. In both cases a subsumption takes place and for that reason these compounds are sometimes called ‘subordinative compounds’.

The representatives of the second subgroup of the appositional compounds are *fighter-bomber* and *girl friend*. They are called ‘additive compounds’. A *fighter-bomber* can be understood as both a *fighter* and a *bomber* at the same time. A *girl friend* is a *girl*, as well as a *friend*. Marchand also states that this is a “semantic distinction of a subjective kind, implying that both elements are equally prominent in the speaker’s mind.” (41) Further on, he points out that additive compounds should not be analysed on the basis of the A + B formula, but rather as ‘B which is also A’. The proper names, such as Austria-Hungary, Baden-Württemberg, can also be regarded as examples of this subclass, as they can be interpreted on the basis of the formula stated above. Finally, it should be mentioned that if compounds of the copula can be analysed on the basis of the formulae AB=B and AB=A, then the relation Subject-Predicate must be reversible. And indeed we can say ‘the girl is a friend’, as well as ‘the friend is a girl’.

The second type of the copulative compounds is represented in h. In this type of copula compounds, we find a copula compound embedded into a more complex structure. Thus, in *doctor-patient gap*, the compound *doctor-patient* is the first member of a combination which has the noun *gap* as its head. In the same manner, a *brother-sister relationship* can be interpreted as a relationship between two people, the one being the brother and the other the sister. Therefore, it can be stated that copulative compounds “demand a reading in which the copula compound is a collective of two members in which each one stands in particular relationship with regard to the following noun”. (Perkles 2008: 50) Thus, the relationship between *doctor* and *patient* and *brother* and *sister* is in coordination with the respective heads. For that reason, these compounds are called ‘coordinative compounds’. Plag (2003: 147) argues that coordinative compounds are in principle ambiguous, since it is possible to derive more than one meaning

from a compound. Thus a *scientist-philosopher crew* can be a crew made of scientist-philosophers, or a crew that is made of scientists and philosophers.

Even though it has been claimed that copula compounds are not very productive in English (Bauer 1983: 203), Olsen (2000c) shows that this subclass of compounds is not marginal. In her study she listed hundreds of attested examples of copulative compounds. Thus, when it comes to the appositional compounds (which she also calls unembedded constructions) (295), she lists a number of semantic groups. She starts with copulatives that denote things and makes a point that there are only a few attested examples of these compounds: *mind-brain*, *tent-office*, *cough-laugh*. By far the most productive semantic pattern is the one that describes people according to their profession, especially to those professions that deal with film, music, media, science, or art. Thus, she gives the following possibilities: film/TV/stage (*writer-director*, *producer-screenwriter*, *director-choreographer*), music (*singer-guitarist*, *dancer-singer*), journalism/writing (*editor-publisher*, *novelist-professor*, *poet-historian*), science (*philosopher-scientist*, *astronomer-author*), art (*artist-designer*, *salesman-artist*), and other (*philosopher-physician*, *explorer-linguist*).

When it comes to the coordinative compounds, Olsen (298) states that there are four general semantic patterns for grouping this type of compounds. These four groups may be generalized into two bigger groups. The first would be the compounds whose elements stand in the 'between' relation to its head. Thus, *doctor-patient gap* would be interpreted as 'a gap between the doctor and the patient'. The second big group are coordinative compounds that have collective nouns as their heads. Collective nouns are grammatically singular elements, whose meaning entails a collection of other elements. Thus, "[t]he type of elements making up the denotation of the collective term can be given expression by the constituents of the copulative, thus, giving a more precise description of the constituents of the whole." (Olsen 2000c: 299) In this manner, a 'duet' made of a 'father and his son' is a *father-son duet*, a 'team' made of a 'man and his wife' is a *man-wife team*, etc.

In connection with the interpretation of the copulative compounds, Plag (2003: 147-148) raises two questions that need to be looked at in a greater detail. The first is whether it is possible to interpret copula compounds as right-headed, even

though at first sight they seem to have two heads. If one labels compounds with regard to their grammatical properties, then it could be argued that copula compounds are right-headed since grammatical markings occur only on the right-hand member of the compound (e.g. There are many *singer-songwriters*/**singers-songwriter*/**singers-songwriters* in this country). Of course, this is a very small evidence for the headedness of copulative compounds, but this still supports the theory that compounds in English are predominantly right-headed. (147) The second question deals with the hierarchical organization and binarity. In the chapter that deals with the main features of compounds, it has been argued that one of the basic properties a compound has is its hierarchical organization. However, as copulative compounds seem to have two equally important elements, it can be argued that this is the case of a non-hierarchical structure. However, if one assumes that copulative compounds are headed, one would automatically arrive at a hierarchical morphological structure. With this proposition, Plag tries to arrive at a “more elegant theory of compounding” (147), however, the whole theory relies on the assumption that one accepts the right-headedness of copulative compounds. The grammatical marking on the right-hand member is one of the points that lead in this direction. However, Plag himself states that “this is only a small piece of evidence for the headedness of copulative compounds”. (147)

4.3 Exocentric compounds

Until now, only compounds that have their semantic head within the morphological unit have been discussed. Therefore, the next chapter will be dedicated to the type of compound that does not have the semantic head. For example, *birdbrain* does not denote the brain of a bird, but rather someone who is not intelligent. Thus, as the semantic head is not to be found within the compound, they are termed ‘exocentric compounds’, with *exo* denoting ‘outside’ and *centric* denoting ‘centre’. Another term for this class of compounds is ‘bahuvrīhi’, a term that originates from the tradition of the ancient Sanskrit grammarians. As far as their morphological shape is concerned, these compounds can be divided into two parts: *bahu-* ‘much’ + *vrīhi-* ‘rice’. Their morphological pattern is X + Noun, where X represents an adjective, noun, or a verb. (Gimborn 2001: 8) However, this pattern does not explain all the examples of exocentric compounds in English (e.g. *showoff* = V +

P), and thus it is necessary to modify it. For a clear analysis, the following morphological pattern (Perkles 2008: 44) will be taken into account:

- (44) a. N + N: baby face, butterfingers, birdbrain, hunchback
 b. Adj + N: redneck, loudmouth, paleface, greybeard
 c. V + N: killjoy, pickpocket, daredevil
 d. V + P: payback, showoff, giveaway, checkup

If one has a closer look at the compounds in (44), one can see that they actually do have something in common with the endocentric compound, namely, that even though they do not have the semantic head, their morphological head is present. *Birdbrain*, *baby face*, or *butterfingers* all belong to the same class as their right-hand member. Therefore it can be argued that “these compounds do have a head and that, at least in terms of their grammatical properties, these seemingly exocentric compounds are in fact endocentric.” (Plag 2003: 146) Unfortunately, this cannot be claimed for all types of exocentric compounds. In order to explain this, not only N + N exocentric compounds should be analysed, but one should also include other types, such as Adj + N, V + N, and V + P.²⁹

As it has already been stated, these compounds are not hyponyms of their heads, and thus from the semantic point of view they are not compounds. However, they do have the morphological head and for that reason can be seen as compounds. As it has been the case with *birdbrain*, compounds such as *redneck*, *paleface*, or *loudmouth* are noun compounds too, because their right-hand element is a compound. When it comes to the semantic exocentricity of these compounds, it seems this type in English is restricted to forms denoting human beings or higher animals (exception is *greenback* as it denotes the US dollar note). (Plag 2003: 146) Additionally, Plag points out that only exocentric compounds of the type Adj + N are moderately productive, whereas the type V + N is extremely rare. Compounds such as *loudmouth* or *greybeard* are sometimes called ‘possessive compounds’, because they denote an entity that is characterized by the property that is expressed by the compound. Thus, a *greybeard* would be someone who has a grey beard, etc. Plag also points out that this type of compounds usually has the Adj + N pattern.

²⁹ This exception is made due to the fact is that it is not possible to give a profound analysis of exocentric compounds only with the example of the quite unproductive N + N type.

Now, let us come back to the headedness of exocentric compounds. It has been stated that they have the morphological head, since their right-hand member decides which word class the whole compound belongs to (e.g. *birdbrain* is a noun because its right-hand member is a noun). However, this cannot be claimed for the examples in d. *Payback*, *showoff*, *giveaway* and *checkup* are neither particles nor verbs, but they are compound nouns. In these cases, we deal with the nominalizations from the respective phrasal verbs, i.e. *to pay back*, *to show off*, *to give away* and *to check up*. Since they are not analysable on the basis AB = B, they are considered to be exocentric compounds. (Perkles 2008: 44) Marchand (1969: 380-389) makes a difference between several types of the V + P compounds. According to him (382) the compounds of the type *showoff* are not neutral in meaning, but they rather have a derogatory meaning. Many of them are slang words, and there are only a few words which are not pejorative (e.g. *go-between* 'intermediary', *standby* 'helper', etc.), which however "do not disprove the general characteristic of the type". (382) In *showoff*, the relation between the elements is one of the subject and predicate. Syntactically, this would mean that the whole compound represents the predicate, while the subject is not overtly expressed. Additionally, similarly to the synthetic compounds, one finds the discrepancy between the structural (*show/off*) and the syntactic-semantic (*showoff/Ø*) analyses. For this reason Marchand (1969) classifies this type of compounds as 'pseudo compounds (13) and defines them as 'combinations with a composite determinant and a zero determinatum" (380).

The second subtype of the V + P formations is exemplified by the remaining compounds in d. Generally, they either denote objects or belong to the predication-type of reference. For instance. A *checkup* means 'general physical examination', or it may refer to any action of checking up. *Payback* is understood as 'return of owned money', while *giveaway* is usually interpreted as 'present'. The third type of these compounds represents mostly those that have *in* as the final element (e.g. *love-in*, *teach-in*, *talk-in*, etc.) They originate from the 1960s, when they were usually used to denote 'group protest'. Later, their meaning changed slightly and has become something like 'group activity'. What is important about this group is the fact that unlike other examples, these compound nouns cannot be considered derivatives from phrasal verbs, because the corresponding phrasal

verbs do not exist (with the exception of *sit in* (at a meeting)). (Marchand 1969: 385) He also argues that the origin of this type is probably sit-down 'sit-down strike'.

Let us now come back to the remaining morphological patterns. It has already been stated that one subtype of the exocentric compounds is called 'possessive compounds', because they denote someone or something that has the features expressed by the both constituents. Plag (2003: 146) argues that this subtype of compounds usually has an adjective as its left-hand constituent. However, considering the examples in (44) a.-c., it can be observed that this is not always the case. Again, within possessive compounds, one can distinguish between two subgroups. (Perkles 2008: 45-46) The first group comprises exocentric compounds of the morphological pattern N + N and Adj + N and they usually denote 'someone who has what is expressed by the two constituents'. Thus, the main function of these compounds is to produce person nouns. Additionally, there are also compounds of this type that denote plants, e.g. *longleaf*, persons, e.g. *popeye*, *greybeard*, or some non-animate things, e.g. *paperback*, *whitecap*, etc. (Marchand 1974d: 334-335) He points out further that semantically, one should analyse possessive compounds as dt + dm/Ø. However, he also notices that the zero morpheme should be seen as "only a classifier, [...] a substantival category[s]er". The zero compound actually converts the Adj + N and N + N combinations from the class 'impersonal' into the class of 'human-denoting'. For that reason Marchand (335) concludes that bahuvrihi compounds are derivatives, and not compounds.

The second type of the possessive compounds is represented in (44) c. Similarly to the rest of the exocentric compounds, the final element in these combinations is not the same as the whole. However, just as the possessive compounds in general, this type of compounds has the semantic head that denotes an agent that is not entailed in the morphological unit. According to Marchand (1969: 380) these combinations "denote the agent who or which performs what is indicated by the predicate/object nexus of the formal basis." Thus, the fact that the noun can be interpreted as the argument of the verb makes them look like synthetic compounds. However, there is one crucial difference between them, namely, the word order. For instance, in *letter writer* the grammatical deep structure is

‘someone writes letters’ (O-P-S). The word order between the constituents is reverse from the one in deep structure. However, in *pickpocket* and *cutpurse* this is not the case. It may be the case that *pickpocket* is ‘someone who (S) picks (P) pockets (O)’ and that *cutpurse* is ‘someone who cuts purses’, but it is evident that the process of lexicalisation “has touched this group of possessive compounds.” (Perkles 2008: 46) These two words do not only refer to a person who only steals purses and whatever he or she finds in other people’s pockets, but it is rather two synonyms of the word *thief*. In the same manner, the remaining compounds in c. a *daredevil*, *killjoy*, and *turncoat* do not refer literally to someone who ‘dares devils’, ‘kills joys’, or ‘turns coats’, but they rather signify ‘a reckless person’, ‘a gloomy person who spoils the fun of the others’ and ‘a traitor’ respectively.

4.4 Meaning relations in English compounds

In the previous chapters each of the most significant groups of nominal compounds have been discussed in detail, without much attention paid to the meaning relations that exist between the elements of compounds. Thus, in this chapter an attempt will be made to clarify the complex relationships that may emerge between the constituents. One of the possibilities to establish the relationships is the theory of argument linking. Plag (2003: 148-152) touches briefly upon this topic. If it is assumed that a given noun-noun compound is in principle ambiguous, the interpretation of the noun will largely depend on the context in which it occurs. Further on, he points out that one should distinguish between at least two different types of nouns - sortal and relational nouns. According to Plag (2003: 148) sortal nouns are those nouns that denote entities, as *chair*, *table*, *window*, etc. In contrast to sortal nouns, relational nouns denote relations between a noun and a second one. Thus, it is impossible for someone to be a mother without being a mother of someone. He continues by saying that “[t]he second, conceptually necessary, entity [...] to which a relational noun relates is called argument.” Thus, if the right-hand member of the compound is a relational noun, the left-hand constituent will be its argument. For instance, *brain surgery* is interpreted as ‘surgery performed on the brain’ and

this process by which a phrase or word from the neighbourhood of a head word is assigned the status of the head’s word argument, is called argument linking. The idea behind this term is that relational

nouns and verbs have empty slots in their semantic representation (the so-called argument structure), which need to be filled by arguments. The empty slots in the argument structure are filled by linking the slots with arguments that are available in the neighbourhood of the noun or verb in question. (Plag 2003: 149)

The process of argument linking may seem straightforward, however when it comes to the primary compounds, it is never possible to know which interpretation one should take into account. Gagné & Spalding argue (2010: 287) “[a]lthough each constituent of an endocentric compound contributes to the overall meaning of the compound [...], the process by which a compound’s meaning is derived is complex and not yet fully understood.” Let us have a look at the following examples: *olive oil*, *baby oil* and *cat rash*. *Olive oil* can be paraphrased as ‘oil made of olives’. However, this relation does not apply to *baby oil*, as this compound would rather be interpreted as ‘oil for babies’. The compound *cat rash* has more than one possible interpretation: ‘a rash on a cat’ or ‘a rash caused by a cat’. As it can be seen, there are compounds that have the same right-hand element (*olive oil*, *baby oil*), but their argument structure is quite different. Similarly, there are compounds that have only one form, but more than one possible interpretation (*cat rash*). The question is, how one can account for these differences and whether it is possible at all to establish these relations between the constituents of compounds. Gagné & Spalding (2010: 287) point out further that even though compounds may have more than one possible interpretation, there is usually one meaning that is intended by the speaker/writer. Whenever there is a situation where the listener/reader has to choose between more than one meaning, he or she will settle for the most likely interpretation in order to understand the utterance in which the compound appears. In other words, in order for the listener/reader to be able to understand the compound, he or she needs other information. And this ‘other information’ is what Gagné & Spalding call conceptual knowledge. The conceptual knowledge is used to construct a relational structure during the interpretation of a compound. Let us have a look at the following example:

- (45) a. This marble museum is built with the most expensive sort of marble.
b. I went to the marble museum yesterday. There is an exhibition about the use of marble in the 19th century.

In the example a. the listener/reader will interpret the compound *marble museum* as a museum built with marble, while in the example b. one will rather interpret the same compound as a museum in which marble objects are exhibited. This factor that is involved in the interpretation of compounds is what Plag (2004: 150) calls the surrounding discourse. Thus, if it happens that the word *marble museum* occurs in a conversation or a text about an exhibition of marble structures, one would probably understand it as ‘a museum where marble objects are exhibited’. In contrast, in a text or a conversation about building materials, the interpretation would rather be ‘a museum that is made of marble’.

Until now, only primary compounds have been dealt with. However, argument linking is also important for compounds whose right-hand element is a noun derived from a verb, namely synthetic compounds. In these compounds, the left-hand element of the compound serves as an argument of the verb. For instances, *beer drinker* is ‘someone who drinks beer’, *letter writer* is ‘someone who writes letters’, etc. Further on, the argument linking theory helps one to analyse synthetic compounds structurally. The compound *letter writer* can be analysed in the following ways:

- (46) a. [[letter write] -er]
 b. [letter [write-er]]

Since **letter write* is not a possible formation, it seems that the example b. is a better solution. After all, a *letter writer* is a writer of letters, which means that the derivative *writer* inherits an empty slot from the verb *write*, and this argument slot can be filled by an *of*-phrase (*a writer of letters*) or by the first member of a compound. (Plag 2003: 149)

However, even though it seems that the meaning interpretation of synthetic compounds is less complicated than in primary ones (it has been stated that the left-hand element serves as the argument of the verb), the argument linking theory sometimes fails. For instance, a *street seller* is not someone who sells streets, but someone who sells something on the street. In the same manner, a *Sunday driver* is not someone who drives Sundays, but someone who drives on Sundays. In these cases “the first element of the compound is semantically not compatible with its possible status as argument” (Plag 2003: 150) and thus an alternative relationship needs to be created. Again, this is where we come to the conceptual

knowledge. Knowing that it is not possible to sell streets or drive Sundays, one will interpret these compounds in the way most likely understood. Thus, it is assumed that every person has a general knowledge of the world, and as it is known that Sundays cannot be driven, this interpretation is automatically excluded from the possible interpretations.

To sum up, if a compound consists of the nouns, meaning of the compound as a whole can be extremely variable. Thus, in order to be able to understand the intended meaning of a compound it is necessary to look at the semantics of the constituents, the possible conceptual relations between the arguments (argument linking), the surrounding discourse, as well as the general knowledge one has about the world.

When it comes to the analysis of coordinative and exocentric compounds, it has to be stated that their analysis is in so far different from the analysis of primary and root compounds, as their constituents are not in the relationship of subordination. For example, *boyfriend* is someone who is a boy and a friend at the same time and thus it can be stated that the meaning of the coordinative compound as a whole can be seen as a sum of the meanings of its constituents. Thus a *physicist-astronomer* is someone who is both a physicist and an astronomer, a *producer-director* is someone who is a producer, as well as director, etc.

Where the meaning relations in compounds are concerned, the most problematic group is represented by exocentric compounds, as their head is not within compound. A *birdbrain* is not a brain of a bird, but a person that is not very intelligent. Thus, it can be seen that the meaning of the whole depends on some other factors that are not represented in the constituents of the compound and that the relation between the constituents is a complex one. Sometimes exocentric compounds denote entities that are metaphorically characterized by the property expressed by the compound. For instance, *loudmouth* is a person that has a 'loud mouth'.

5. Composition in Serbian

As Babić (1986: 30) states, composition is that kind of process where one word is created from two or more words or two or more bases³⁰. The product of this process is called a composite word. Since this term very often overlaps with the term compound, and sometimes is even used as a synonym for compound³¹, an attempt to delimitate them must be made. In order to avoid this confusion, a short overview of the types of composition in Serbian will be given. For this purpose, Babić's *Word formation* will be the main source and his classification will mainly be taken over. However, it has to be stated that types of composition often overlap and thus it is necessary to give a short overview of different types in order to be able to distinguish compounding from other types of word-formation. The most important ones are: pure compounding, suffixal compounding, coalescences ('sraslice'), half-compounds and prefixation.

Pure compounds are those words whose second part is an independent word, such as *bratoubistvo* ('murdering of a brother') *minobacač* ('mortar'), etc. (Babić 1986: 30). Pure compounds with the binding vowel *-u-* are limited only to some words that have the vowel in the first part: *polukrug* ('semicircle'), *polumrtav* ('half-dead'), etc. He also gives examples of the pure compounds with the binding vowel *-i-*, as in *kažiprst*. When it comes to suffixal compounding, these examples are more common compared to pure compounds. These compounds are composite words that after the process of composition also received a suffix, as in: *čudo* ('miracle') + *tvoriti* ('create') + *-(a)c* > *čudotvorac* ('miracle worker'). Since this type of compounds is a very productive one, and many of them have the Noun + Noun form, they will be discussed in detail later in this paper.

The next group to be discussed are coalescences. This type of composition is probably the most controversial one, as different authors classify them in different ways. Therefore it is necessary to devote some space to it in order to be able to make a distinction between compounds and coalescences. Stevanović (1964) does not make any differences between them and he defines coalescences and compounds in the same way, as words that consist of two or more words (414)

³⁰ He also gives examples of compounds with three bases. However, he claims that these are extremely rare.

³¹ See Stanojčić (2004: 133)

and he analyses in the same way *dangubiti* ('to waste time') from *dan* ('day') + *gubiti* ('lose'), *akobogda* ('hopefully') from *ako* ('if') + *bog* ('god') + *da* ('gives') and *jugoistok* ('southeast') from *jug* ('south') + *o* + *istok* ('east'), *plavokosa* ('a girl with blond hair') from *plav* ('blond') + *o* + *kosa* ('hair'), etc. Babić (1986: 31) states that coalescences are words that originate from a constant group of words that have developed into a complex word, i.e. *blagdan* ('religious holiday') from *blag* ('gentle') + *dan* ('day'). He says that this process is called coalescence and products of it are coalescences. This is not a frequent way of creating new words, and a lot of them are sometimes not even products of word-formation, i.e. the city in the western part of Bosnia is sometimes called *Banja Luka*, gen. *Banje Luke*, but today it is more common to hear *Banjaluka*, gen. *Banjaluke*. He also states that coalescences can also be words without a binding vowel, such as *duvankesa* ('tobacco pouch') from *duvan* ('tobacco') + *kesa* ('pouch').

Klajn (2002: 28) argues that coalescences can be distinguished from real compounds on a clearer and safer basis. The presence or the absence of the binding vowel should not be the only criterion for making this distinction, but one should take into account the syntactic relationship between the words involved in this process. He points out that the term coalescences should be applied to those words that can be used as syntagmas in the same form and the same order. According to this *Beograd*, *dangubiti*, etc. are coalescences. In contrast to Babić (1986: 31) who claims that words such as *duvankesa* ('tobacco pouch'), *zimzelen* ('evergreen'), or *Ivangrad* are examples of coalescences, Klajn (28) argues that this cannot be the case, since these words could never function as two words. He defines them as compounds without binding vowels. Further on, he points out that words such as *spomen-dan*, *tempera-boje*, *radijus-vektor*, etc. are not coalescences either, but should rather be considered as half-compounds. They are not syntagmas and they could not appear in a sentence as two consecutive words. It can be considered that compared to other types of compounds, half-compounds have a specific feature. Namely, both parts of the word retain their accent and they are usually marked by a hyphen between the two constituents. However, Klajn also points out that it would be wrong to consider every word-formation that has a hyphen in it as a half-compound, since spelling cannot be a criterion for categorization in word-formation.

The next type of composition, namely prefixation, is an interesting one, as it is used for two different types of word-formation in English and Serbian. As Marchand (1969: 129) states, prefixes “are bound morphemes which are preposed to free morphemes” and thus the process of adding a prefix in front of a free morpheme is called prefixation. However, Babić (1986: 33) claims that prefixation includes those processes of word-formation in which a free morpheme is combined with a preposition, as in *predstraža* (‘advance guard’) the negation particle *ne*, as in *nečovjek* (‘mean person’), or with suffixes that are not attested as independent words, as in *popiti* (‘drink out’)³². In other words, the confusion comes from the fact that according to Marchand, the combinations Preposition + free morpheme cannot be considered as products of prefixation, as prepositions are independent words and prefixes are not. For this reason, Babić’s proposal to analyse *predstarža* and *popiti* in the same manner is not justified and in this paper strict delimitation between prefixation and compounding is going to be made.

Finally, there is a type of composition where two words are joined together in such a way that only some grammatical features are lost, while each of the words retains its accent, and mostly its meaning. These words are called half-compounds and they are marked by a hyphen between the left-hand and the right-hand constituent, i.e. *radio-stanica* (‘radio station’), *rak-rana* (‘serious wound’), etc. Interestingly, Babić (1986: 32) claims that compounds that have *aero-*, *auto-*, *foto-*, etc. as one of its constituents should be considered as half-compounds. However, as it has already been discussed in chapters about English compounds, these bases have their origins in neoclassical languages (Greek and Latin) and thus they will be termed neoclassical compounds. Half-compounds were frequently used in poetical texts in the 19th century and in these texts it is possible to find half-compounds that are used only for poetical purposes, and thus they are called poetical half-compounds.³³ (Babić 1986: 32) Therefore half-compounds can be considered as a language phenomenon that is on the border between syntax and word-formation, because it is assumed that in the process of word-formation a new unique word is created and as it has been seen constituents in half-compounds mostly keep their original meaning. However, they will be a matter of discussion later in this paper.

³² For different opinion see Klajn (2002)

³³ See Babić (1986: 32) for more information about half-compounds

5.1 Prefixation and compounding in the Serbian literature about word-formation

The notion of prefixes and prefixation has appeared quite lately in Serbian linguistic literature. (Klajn 2002: 173) For instance, Stevanović (1964: 443-471) in his comprehensive book wrote one chapter “Compounds with prefixes” in which at the very beginning he defines prefixes as prepositions that become prefixes when they are combined with the other part of the compound. Further on, he points out that when it comes to the analysis of compounds with prefixes one should take into account only those compounds where a prefix is combined with another word that has not undergone any formal changes or changes in meaning. In other words, it means that the other word does not have any affixes. However, Klajn (2002: 173) argues that according to this selection, Stevanović excludes all the examples of the so called prefixal-suffixal word-formation. The reason for this exclusion is not mentioned. According to Klajn it is possible to assume that Stevanović did this because he was planning to handle these examples under corresponding suffixes. Babić (1986: 33) acknowledges that prefixes are not independent words and thus should not be a matter of discussion in relation to compounds. However, then he states that because of the uniqueness of this type of word-formation, they will be considered as a certain type of compounds and not as a product of derivation. Nevertheless, in the chapters devoted to adjectives, nouns, verbs and adverbs, he always handles prefixal verb-formation under a separate title and not within compounds. Stanojčić & Popović (2004: 133) state that one very productive way of creating new compounds is by prefixation, which means that they completely incorporated prefixation within compounding.

According to Klajn (2002: 175) the reason for this confusion lies in the fact that none of the previously mentioned grammars try to define prefixes. Most of the authors rely on the previous literature about prefixes, mostly on the works by Maretić (1899). He does not recognize the fact that prefixes as *raz-*, *pro-*, etc. are not used as prepositions any more and thus are a subject of prefixation. However, as his work is based on the diachronic point of view, he was always able to find examples where these words were used as independent words – as prepositions. The later grammarians introduced the term prefix, but they still treated it as a preposition that in combination with a noun, verb or adjective can form a

compound. What Klajn (2002: 178) suggests is that prefixes should be considered as affixes, since they are not independent words, and therefore prefixation cannot be a matter of discussion within compounding. The only question that should be answered is whether prefixation should be included in derivation or if it should be acknowledged as a separate category of word-formation. Klajn's solution, namely the one that states that prefixation and suffixation are seen as categories of a more general term, namely affixation, will be taken over in this paper. Therefore, prefixation is going to be separated from compounding, and thus is not going to be a matter of discussion.

5.2 What is a compound in Serbian?

As it has already been stated in the chapter about English compounds, it is not possible to give a universal and the only true definition of compounds. Thus, the definition and description of compounds in Serbian is controversial in the same way. Stevanović (1964: 414) defines compounds as words that consist of two or more words and says that they actually originate from certain sentence parts or whole sentences that are composed of at least two words.

- (47) a. dangubiti ('to waste time') from dan ('day') + gubiti ('lose')
b. jugoistok ('southeast') from jug ('south') + o + istok ('east')

As it has been the case in the chapters about English compounds, it is also in Serbian highly problematic to reduce compounds to two words, because we have already seen that there are instances where it is not possible to state if one of the constituents of the compound is actually a word. Again, the most problematic cases are neoclassical compounds. *Biosfera* ('biosphere') can without any doubt be separated into two units *bio* (Greek 'life') + *sfera* ('sphere'). *Sfera* is attested as an independent word. The same cannot be said for *bio*. Thus, the problem with Stevanović's theory is the fact that some important aspects of compounding have been omitted, in so far as he does not deal with neoclassical compounds at all. Apart from that, he mainly concentrates on the general theory about compounds, i.e. the change of meaning in compounds, stress, alternative forms, etc. Furthermore, a discussion about formal criteria of compounds is omitted and one very important question is put aside, namely which forms can be considered as compounds. This has been a very controversial topic among Serbian linguists, as

until today they have not agreed on what one should consider as compounds. (Barić 1980: 15)

Barić (1980: 15) argues that compounds are those words that have been motivated by two words. This can be seen as one of the criteria, however not as the only one. As already stated, it is not possible to reduce compounds to two words. In Serbian, examples of compounds where one can find whole words as its constituents are extremely rare (apart from coalescences and half-compounds). Babić (1980: 30) argues that compounds are those words that are created on the bases of two or more bases or of two or more words. Klajn (2002: 22-23) also accepts this point and adds that until today a better solution has not been found. Klajn also states that Barić's criterion of motivation as the only criterion cannot be accepted, because she does not mention the content of the compound at all. When it comes to spelling as a criterion for defining compounds, one can argue that in some cases it can be helpful. For instance, in examples *soda-voda* ('soda water') and *duvankesa* ('tobacco pouch') the spelling would be enough to say what a compound is. As it has already been stated, all half-compounds in Serbian have a hyphen between the two constituents. However, as all word-formation products in Serbian are written together or with a hyphen, other criteria are necessary to distinguish compounding from, for instance prefixation or suffixation. For example, *parobrod* ('steamboat') and *nadjačati* ('overpower') belong to compounding and prefixation respectively. However, the spelling only could not tell us the distinction. For this reason, in order to isolate compounds from other types of word-formation, just as with English compounds, it is necessary to have a look at morphological, syntactic and semantic properties of the word. Again, similarly to English, different approaches (morphological or syntactic-semantic) will result in a different classification of compounds. But before we start discussing them, it is necessary to have a closer look at their morphological and syntactic-semantics properties.

6. Classification of compounds in Serbian

6.1 Morphological and syntactic-semantic properties of N-N compounds in Serbian

When it comes to the morphological features of N-N compounds in Serbian, it is necessary to have a look at all elements that one compound is created of. Thus, Stevanović (1964: 414) argues that compounds are complex words composed of two or more words, assuming that words only are involved in this process. Babić (1987: 30) expands the definition and states that compounds are not only built of two or more words, but also of two or more bases. Finally, Klajn (2002: 22) states that in compounds, the first part can be a base or a word, while the second part must be a word, as in *poljo-privreda* ('agriculture') or *ruko-voditi* ('to run, to operate'), etc. However, Rakić (2007: 771) argues that Klajn's proposition cannot be accepted for two reasons. Firstly, in examples *palvook* ('blue-eyed') or *crnokos* ('black-haired'), the second constituent is definitely not a word. *-ok* is derived from *oko* ('eye'), while *-kos* comes from *kosa* ('hair'). As Rakić argues, Klajn tries to justify his theory by saying that *plavook* and *crnokos* are adjectives, because a zero suffix is added at the basis of the words *oko* and *kosa*. According to Rakić, the problem here is that Klajn does not actually define the zero suffix and on which bases it can be added. Thus, he proposes that a solution for examples like these, one should take conversion into account. Let us have a look at the following example:

(48) *devojka crne kose* ('a girl with black (dark) hair')

What Rakić proposes is that the noun phrase *crna kosa* is converted into the compound *crnokos* and this transformation can partially be motivated by the adjectival use of the noun phrase as in the example above, where *crne kose* obviously has the adjectival function. The situation becomes more complicated with the examples of adjectival and noun compound that have the same structure:

- (49) a. *crnorep* (adj.) ('having a black tail') < *crn* ('black') + *o* + *rep* ('tail') vs.
 crnorep (n.) ('something having a black tail') < *crn* + *o* + *rep* + *Ø
 b. *trolist* (adj.) ('trefoil') < *tr-* ('three') + *o* + *list* ('foil') vs. *trolist* (n.) ('something that is trefoiled') < *tr-* + *o* + *list* + *Ø

Obviously, it does not make enough sense to claim that in one case the adjectival zero suffix is added, and in the others the nominal. Thus, Rakić proposes that it is

unnecessary to introduce the term conversion in the definition of the morphological structure of compounds and derive, for instance, nouns from adjectives. Instead, it is enough to say that the second constituent of a compound can be a root base, and the compounds of the type *crnokos* should simply be defined as compounds whose head is the left-hand constituent.

The other problem with Klajn's definition is the description of the morphological structure of the left-hand constituent. Klajn states that the first part of the compound can either be a word or a root base. However, he does not say whether, for instance, a word can include prefixes and suffixes or only flexional endings. Klajn (2002: 22) gives examples, such as, *očenaš* (the Lord's Prayer) < *oče* (vocative 'father') + *naš* ('our') in which the first element *oče* is the vocative form of *otac*.³⁴ Thus, Rakić argues that it should be examined whether a compound can have affixes, or even compounds as its first element. According to Klajn's list, there are very few examples where the compounds have a compound as the first part and those are mostly numbers or names, such as:

- (50) a. *dvadesetogodišnjica* ('the 20th anniversary') < [[*dva* + *deset*] + *o* + *godišnjica*]
b. *Beogradput* ('the Belgrade roads') < [[*beo* + *grad*] + *put*]

Cases where the first part of the compound is a derivative are also extremely rare, and according to Rakić, most of the examples found are those compounds that have an adjective as its first part, such as *Južnoamerikanac* ('South American'), *smaragdnazelen* ('emerald green'), etc. Additionally, Rakić asks the question whether it is legitimate to include these examples in the general definition of compounds and therefore proposes a further analysis. The first part of the compounds is in most cases a morphologically simple word. However, there must be given an insight into other possible structures. One of the first things to do is to clarify whether the examples *Južnoamerikanac* and *smaragdnazelen* are compounds or half-compounds. One of the most important aspects that has to be taken into account is accent. Even though Babić (1986) claims that these words have only one, they are usually pronounced with two accents. (Rakić 2007: 772) He also points out that in favour of this theory also speaks Klajn's observation that formation of the type *kulturnoistorijski* ('cultural historical'), *naučnoistaživački*

³⁴ For sound alternations in *otac* (nom.) – *oče* (voc.) see Stanojčić & Popović (2004: 58)

(‘scientific explorative), *društvenopolitički* (‘social political’) can be seen as bi-adjectival, in which case they would be written with a hyphen: *kulturno-istorijski*, etc. and that is the way of writing that Klajn connects with the separate accent of each of the compound members. Thus, they must be classified as half-compounds. Therefore, Rakić concludes that in these cases compounds have complex first parts, and in most cases those are adjectives *južno*, *smaragdno*, *otvoreno*, *grupno*, etc. and the whole combinations are mostly adjectival compounds. The next step is to answer the question if there are N + N compounds that have prefixal or suffixal derivatives as their first element. The only examples where the first part is a complex words are suffixal derivatives, as in *Srbijaprevoz*, *daljinometar* and *Slavonijapromet*. However, as these words are pronounced with two accents (Rakić: 2006) they are considered to be half-compounds. Thus, Rakić (2007: 772) concludes that N + N compounds in which the first part is a derivative are half-compounds in which each of the elements has its own accent. Therefore, he proposes the following possible morphological structure of N + N compounds:

- (51) a. both elements are morphologically simple words: N + N, as in
 vjeroispovijest (‘religion’) < *vjer-* (‘religion’) + *o* + *ispovijest* (‘confession’)
 b. the second element is a word derived by conversion: N + N conv.
 c. the second element is a word with an affix: N + (X + suffix) or N + (Prefix + N)

It has to be mentioned that nouns in compounds are not whole words in most cases, but noun bases (as in *vjeroispovijest*) and it also has to be stated that this classification will be adopted in this paper.

The next step is to have a closer look at the syntactic-semantic properties of Serbian compounds. As far as this aspect of analysis is concerned, it is necessary to look at the relationship between the constituents of the compound. As Klajn (2002: 33) argues, syntactic relationships between them can be extremely complex and therefore none of the classifications and analyses have generally been accepted until today. He argues further that when it comes to the syntactic analysis of compounds, one should always look for the presence or the absence of the copula. According to this, the presence of the copula would assume that the relationship between the constituents is one of coordination, while the absence of it would mean that constituents have a subordinative relationship. Let us have a look at the following examples:

- (52) a. gluvonem ('deaf and dumb') < gluv ('deaf') + o + nijem ('dumb')
b. duvankesa ('tobacco pouch') < duvan ('tobacco') + kesa ('pouch')

As can be seen, the constituents under a. have a coordinative relationship, as the copula 'is' can be inserted and the word order can be changed without altering the meaning of the compound. In contrast to this, the constituents under b. are in the relationship of subordination, as it is not possible to insert the copula. Actually, Klajn (36) argues that these compounds are derived from syntagmas³⁵ and thus the relationship between the constituents is unequal (subordinative)³⁶. Additionally, since this subordination between constituents originates from syntagma constructions, one can assume that constituents can have many different relationships, as determination, modification, subordination, etc. and so it is clear that the border between different types of compounds will be blurred. Therefore, only when one takes into consideration all these different aspects, will it be possible to give an acceptable classification and description of N+N compounds in Serbian. Thus the following task is to look at each of categories separately and analyse them in detail.

When it comes to the internal structure of compounds, it is necessary to analyse their morphological, as well as syntactic-semantic features. As Klajn (2002: 33) argues, since syntactic relationships between compound elements are extremely complex, it would not be possible to give a universal categorisation. However, looking at the works of Stevanović (1964), Babić (1986) and Klajn (2002) one could argue that in the case of syntactic analysis of compounds, there are three basic categories or types: copulative (coordinative), determinative (modifying) and rectional (complement) compounds. However, only the syntactic analysis is not enough to give a comprehensive description of noun compounds and thus it is necessary to look at their semantic features as well. In connection with this, the notions of endocentric and exocentric compounds will be examined. As it is the case in English, this classification is done on the basis of the presence or the absence of the head of the compound. Therefore, the final classification of compounds will be similar to English. It is going to be a mixture of syntactic and semantic features that results in three major categories - endocentric, exocentric

³⁵ For similar opinion about English compounds see Marchand (1967)

³⁶ For more information about coordination and subordination see chapters Determinative compounds and Coordinative compounds

and copula compounds, taking into account that some of them will be divided into smaller subgroups.

6.2 Endocentric compounds

Endocentric compounds are those compounds whose semantic head is within the compound. Just as in English, Serbian endocentric compounds can be represented by the formula $AB=B$, where A is the determinant and B is the head or the determinatum.³⁷ For example, in *parobrod* ('steamboat'), the word *brod* ('ship') is determined by the word *para* ('steam') and thus is the head of the compound. It is important to mention that the compound retains all the grammatical features that the head originally had. For instance, the genitive form of *brod* is *brod-a*, and thus of *parobrod* is *parobrod-a*, instrumental *brod-om* and *parobrod-om*, etc. Thus, every compound that can be analysed on the principle $AB=B$ will be considered as endocentric. Stevanović (1964: 425-428) and Klajn (2002: 33-36) distinguish between two types of endocentric compounds: determinative (Serb. *odredbene*) and complement (Serb. *dopunske*) compounds. Syntactically, both types belong to one more general group, namely subordinative compounds, as the syntactic relationship between the compound constituents is one of subordination.

6.2.1 Determinative compounds

According to Stevanović (1964: 426) "determinativne, odredbene su složenice postale od sintagmi u kojih je je jedan dio određivao drugi [determinative compounds are those compounds that originate from syntagmas in which one part determines the other]". However, as Klajn (2002: 33) argues, it is not directly stated in which direction this determination goes; however, it is to be assumed that in most cases the left-hand part determines the right-hand one.

Let us start with the nominal compounds without affixes. According to Barić (1980: 27) the process of creating a new determinative compound without any process of prefixation or suffixation involved, looks as follows:

1. from two words (*bacač* ['thrower'] and *mina* [Gen. Pl. 'mine']) one word is created (*minobacač* ['mortar'])

³⁷ Marchand's (1969: 11) definition of endocentric compounds is also going to be applied to compounds in Serbian

2. two accents (*bàcač* and *mina*) are reduced to one (*minobàcač*)³⁸

3. from two meanings (*bacač* and *mina*), one unique meaning is derived (*minobacač*)

As Barić (28) points out further, it is important to mention that the relation between the elements within the compound is different from the relationship that the constituents used to have before they were involved into the process of compounding. This different type of relationship between the constituents within a compound is reflected on several levels: on the morphological level, the new created compound undergoes only one declination, namely that of the right-hand element, i.e. Gen. *minobacač-a*; on the syntactic level, the compound has one attribute, and not two for each of its elements, i.e. *stari minobacač* ('old mortar'); on the semantic level, the compound is one lexical unit. Thus, it follows that the newly created compound is not a mere product of the meanings of the two constituents, but it is a new dimension of those two meanings. For the reasons of getting a deeper insight into the structure of nominal compounds, it is necessary to give a detailed description of their morphological and syntactic-semantic properties.

The first thing to start with is the word-formation analysis. Nominal compounds of non-prefixal and non suffixal word-formation can be divided into several word-formation types, based on the presence or the absence of the binding vowel in the compound. Thus, according to Barić (1980: 39) three categories can be distinguished:

1. The word-formation type with a binding vowel, i.e. *tekst* + *o* + *pisac* ('text writer')
2. The word-formation type with the binding vowel -Ø, i.e. *amper* + Ø + *metar* ('ammeter')
3. The word-formation type with a lexical morpheme that ends in -o, which however cannot be separated from the lexical morpheme in the word-formation analysis. Thus, those compounds are without binding vowel, i.e. *hidro* + *elektrana* ('hydroelectric power station')

³⁸ For more information about accents in Serbian see Stanojčić & Popović (2004)

Before turning to other points, it is necessary to give several comments on Barić's classification. First of all, in the chapter about binding vowels it has been stated that something like a zero binding vowel does not exist, and thus the second category should be renamed into compounds without binding vowel. Secondly, the third category comprises neoclassical compounds (compounds with Greek and Latin elements in it) and therefore they will not be discussed in this paper in a greater detail, as their status in Serbian is as controversial as in English, since it is not possible to classify them either as lexemes or as affixes. Apart from the classification according to the binding vowel, it is possible to divide compounds according to the elements that a compound consists of. The following categories have been proposed by Klajn (2002: 41-51) and Barić (1980: 39):

1. compounds that consist of a noun base and a noun, i.e. *vjer-* + *o* + *ispovijest* ('religion')
2. compounds that consist of two nouns, i.e. *čuvar* + *kuća* ('something that keeps a house safe')
3. compounds that have German or Latin elements as one of its constituents, i.e. *hidro* + *elektrana* ('hydroelectric power station')

6.2.1.1 The type noun base + binding vowel + noun

As already stated, compounds that have words of Greek or Latin origin will not be dealt with in this paper and thus the third group will be left out. Let us start with the compounds that consist of a noun base and a noun. According to Klajn (2002: 41), most compounds of this type are either determinative or rectional. As it will become clear in the next chapter, it is not very helpful to classify compounds to determinative and rectional, as it is very often the case that it is not possible to draw a clear border between these two classes. For instance, *vjeronauka* ('religious instruction') can be analysed as 'nauka o vjeri' ('science about religion') or as 'verska nauka' ('religious science'). Klajn (42) argues that this group originally was unproductive. However, this expanded largely thanks to new compounds such as *parobrod* ('steamboat'), *drvored* ('row of trees'), *vjeroučitelj* ('religious teacher'), *prestolonasljednik* ('the crown prince'), etc. Many of these compounds are actually translations from other languages, especially from German, such as *parobrod* (*par* + *o* + *brod*) from *Dampfschiff*, *redoslijed* (*red* + *o* +

slijed) from *Reihenfolge*, *djelokrug* (*djel* + *o* + *krug*) from *Wirkungskreis*, *strahovlada* (*strah* + *o* + *vlada*) from *Schreckensherrschaft*, etc. *Nosorog* (*nos* + *o* + *rog*) is a special example as it is a translation of the German word *Nashorn*, and the German term is the translation of the international Grecism *rhinoceros*. This example is a special one for one more reason, in so far as this compound actually should not be classified as a determinative, since its semantic head is not within the compound and the meaning of the whole word has to be reconstructed on the basis of the constituents of the compound. Compounds of the type *bratoubistvo* ('fratricide'), *djecoubistvo* ('killing of children'), etc. probably have their origins in the Latin words that end in *-cidium*, as *parricidium*, *fatricidium*, either directly or indirectly via German words *Vatermord*, *Brudermord*, etc.

As it has been seen in the previous examples, when it comes to the binding vowel, in most cases it is *o*. (Klajn 2002: 43; Barić 1980: 45) The binding vowel *e* can be found in the following examples: *oceubistvo* ('patricide'), *kraljeubistvo* ('regicide'), and all the other examples with the noun *ubica* as its second element.

6.2.1.2 The type noun + noun (without binding vowels)

The majority of compounds of this type are loan words from Turkish or German. When it comes to Turkish loans, most of the compounds are archaic today, or the meaning of at least one constituent is unknown to the speaker of the Serbian language today. According to Klajn (2002: 45) their status is rather controversial, as it is not clear whether they should be analysed as compounds or half-compounds, i.e. *đul-baklava* ('a sort of baklava'). As already stated, it is very often the case that meaning of the one of the constituents is unknown to the native Serbian speaker, and in this case it is *đul*. As far as compounds of German origin are concerned, they are much more common than Turkish loans. Among German loans, there is a great amount of half-motivated compounds, as *flispapir*, *lajtmotiv* ('leitmotiv'), *maskenbal* ('masked ball'), etc. where the second part is a Germanism or an internationalism, while the first part is always the word that appears only in this combination.

Klajn (45) also points out that there are opposite cases where the first part is known in Serbian, while the second part is unmotivated, i.e. *šrafčiger* ('screwdriver'), *kulturtreger* ('bearer of culture'), etc. A very interesting case is the

pseudo-compound *šofer-šajbna* ('windshield'), which according to Klajn (46) "pokazuje da se nemački tvorbeni obrazac može preuzeti i bez preuzimanja obeju reči iz originala [shows that the German word-formation pattern can be taken over without actually taking over the two words from the original language]". The German word for 'windshield' is *Windschutzscheibe* and thus it is very probable that the first part has been replaced by the more familiar word, namely *šofer*. When it comes to compounds without binding vowels that are of native origin, it has to be stated that this pattern is extremely unproductive (Klajn 2002: 48; Barić 1980: 76). There is one example that is usually given as the representative of this subgroup, namely *čuvarkuća* ('common houseleek' ('sempervivum tectorum')), from *čuvar* ('guard') + *kuća* ('house'), which is an imperative compound derived from 'čuvaj kuću' ('guard the house').

6.2.1.3 Compounds derived by affixes (composite-suffixal word-formation)

Compounds derived by this type of word-formation pattern are the most common in Serbian (Babić 1986: 62; Klajn 2002: 53) There are several morphological patterns according to which these compounds are created:

1. noun base + noun base + suffix
2. nouns base + adjective base + suffix
3. noun base + verb base + suffix

At first sight it might seem that the second and the third group actually do not fit the topic of the paper, as the main subject is compounds that consist of two nouns. However, as it has already been agreed, compounds are binary constructions, which means that they are made of two elements. Thus, the second element in adjective base + suffix and the verb base + suffix respectively, which means that by adding the suffix, a noun is created and therefore they fit the analysis of noun + noun compounds. Further more, it has to be mentioned that some authors handle these compounds as another word-formation type, namely affixation "jer je izvođenje [...] osnovniji način nego slaganje [because affixation is a more common way of creating new words than composition]." (Babić 1986: 62) However, as it has been the case with English synthetic compounds, we will consider this type of compounds in Serbian as products of compounding and not as products of

affixation. Thus, the two possibilities, compounding inside derivation and derivation inside compounding can be represented in the following way:

- (53) a. [[base + base] + affix]
b. [base + [base + affix]]

The type noun *base* + noun *base* + *affix* is typical of zoological terminology. In most cases these compounds denote a whole class and therefore usually appear in the plural form, i.e. *bodljokošci* ('echinodermata') from *bodlj-* + *o* + *kož-* + *-ci*³⁹, *mrežokrilici* from *mrež-* + *o* + *krill-* + *-ci*, etc. The only example that is used widely in the singular form is *vodozemac* ('amphibians') from *vod-* + *o* + *zem-* + *-ac*. This example is also special for one reason. Namely, the noun base of the second part of the compound is *zemlj-*, however probably because of the phonological reasons the *lj* has been omitted, and this process is called shortening of bases. (Klajn 2002: 51, Babić 1986: 62) There are also examples with the suffix *-je*, as *rudogorje* ('mountain rich with minerals') from *rud-* + *o* + *gor-* + *-je*, as well as those with the zero (\emptyset) suffix, as *vukodlak* ('werewolf') from *vuk* + *o* + *dlak* + \emptyset .⁴⁰

As was the case with the previous group, the type noun *base* + adjective *base* + *suffix* is not very productive. According to Klajn (2002: 52) there are only few examples that without any doubt fit into this group: *zemljouz* ('isthmus') from *zemlj-* ('land, soil') + *o* + *uz-* ('narrow') + \emptyset (*uz-* comes from *uzak*), *moreuz* ('straits') from *more-* ('sea') + *uz-* ('narrow') + \emptyset , and *zimzelen* ('evergreen plants') from *zim-* ('winter') + *zelen-* ('green') + \emptyset . Apart from these examples, Klajn mentions a few other compounds whose status as compounds is sure.

Now, let us turn to what according to Klajn (2002: 53) and Babić (1986: 62) is the most productive way of creating nominal compounds in Serbian, namely noun *base* + verb *base* + *suffix*. At the very beginning it is important to point out that at first sight the verb base sometimes has the same form as a deverbal noun that ends with a consonant, however they must not be regarded as identical. For instance, *-vod* in *naftovod* ('oil pipeline') and *-rez* in *drvorez* ('wood carving') are obviously not the same words as nouns *vod* and *rez*. In defining these words, one would not use a noun but rather a verb. Thus, *naftovod* is 'uređaj za dovođenje nafte' ('a device for delivering oil') and not 'dovođenje nafte' ('delivering of oil') and

³⁹ *ž* becomes voiceless in front of the voiceless *c* (assimilation according to manner of articulation). For further information, see Stanojčić & Popović (2004: 45)

⁴⁰ This example will be dealt with in a greater detail in the chapter about exocentric compounds

drvorez is rather interpreted as ‘rad izrezan na drvenoj ploči’ (‘something that is carved into a wood plate’) and not as ‘rezanje drva’ (‘carving of woods’). Klajn (2002: 53) also points out that there are border cases, as for instance, *ribolov* (‘fishing’) can be interpreted either as ‘lov na ribu’ (‘hunting of fish’) or as ‘loviti ribu’ (‘to hunt fish’).

The subgroup with the morphological pattern *noun base + verb base + -ac* is a very homogeneous one. With the exception of the compounds *kitolovac* (‘whaling ship’), *minolovac* (‘minesweeper’) and *ledolamac* (‘icebraker’), all the other compounds are *nomina agentis*, which means that they denote human beings, and the noun is almost in every single case the object of the verb. (Klajn 2002: 56) This difference in the function of the noun bring us to another problem in connection with the classification of the compounds in Serbian, since some linguists classify compounds according to the relationship that the constituents used to have in the syntagma before they were turned into a compound. As has been stated previously, apart from determinative compounds, there is another group according to Stevanović (1964: 428-432) that is called rectional or complement compounds. However, as it will become clear in the next chapter, a clear line between determinative and rectional compounds cannot be drawn as they overlap in many cases.

6.2.2 Rectional (complement) compounds

Rectional (complement) compounds are defined as those that “postale od delova koji su jedan s drugim stajali u zavisnom odnosu kao upravni član i njegova dopuna – najčešće od nekog prelaznog glagola i njegova objekta [that originate from elements that used to stay in a dependent relationship as the determinative part and its complement – usually from a transitive verb and its object]”. As Klajn (2002: 35) argues, as it is sometimes extremely difficult to differentiate between a determinant and a complement in a sentence, it is logical that this category in many ways overlaps with the previous one, namely the determinative compounds. For this reason Stevanović (1964: 428) points out that only those compounds that originate from a verb and its objects are undoubtedly rectional. He exemplifies it with the following examples of compound nouns: *pismonoša* (‘mail carrier’) from

‘nositi pisma’ (‘carry letters’), *vlastoljublje* (‘craving for power’) from ‘ljubiti vlast’⁴¹ (‘to desire power’), etc. In the cases where the complement is not in the accusative form, but in the instrumental one, Stevanović argues that these compounds have more determinative than rectional character. Therefore Klajn (2002: 35) argues that it is more natural to analyse the noun *parobrod* (‘steamboat’) as ‘parni brod’ (‘steam boat’) than ‘parom pokretan brod’ (‘by steam operated boat’). In many other examples, such as *ljubomora* (‘jealousy’), from ‘ljubav mori’ (‘love aches’) and *umotvorina* (‘creation’) from ‘um tvori’ (‘mind creates’) it is not clear in which case the noun is from the first part. Further more, Klajn points out that there is another problem that Stevanović does not recognize, in this case it is the status of the so called imperative compounds. The main feature of these groups is that the verb always stands in front of the noun, whereas the noun is in most cases object, and rarely subject. The following examples are given: the noun is object, i.e. *gulikoža* (‘usurer’) from ‘guli kožu’ (imper. ‘tear the skin!’); the noun is subject, i.e. *visibaba* (‘snowdrop’) from ‘visi baba’ (‘hang, granny!’). Further on, he points out that when it comes to those imperative compounds where the noun is the subject, one could definitely talk about rectional compounds. However, when it comes to the those with the subject, Stevanović (429) argues that “prvi, glagolski dio determinativni, a drugi, imenički – nekadašnja je upravna reč [the first, verbal part is the determinative part, and the second is the modifying one]” and this is what Klajn (2002: 35) considers to be the biggest problem with the imperative compounds. Namely, the group with the subject cannot be distinguished from determinative compounds, since the compounds from this group have the determinant and the modifier.

Since the border between determinative and rectional compounds is extremely blurred, Klajn (2002: 36) proposes that this classification should be abandoned. The core of the problem lies in the fact that there are too many relationships between words within a syntagma (i.e. determination, modification, cohesion, etc.) and thus it is very improbable that compounds that originate from those kinds of relationships can be classified in a straightforward manner. Therefore, he proposes that a better solution would be to classify compounds according to the quality of the relationship between compound members. Thus, as in determinative

⁴¹ In this case the verb *ljubiti* (‘kiss’) is used in the old-fashioned manner in the meaning ‘to love, to desire’

and rectional compounds the relationship between constituents is not equal, they would be united into one larger group - subordinative compounds. In case none of the elements are subordinated to each other, i.e. *gluvonijem* ('deaf and dumb') from *gluv* ('deaf') + *o* + *nijem* ('dumb'), one should classify them as the so called coordinative compounds. Thus the next chapter will deal with the compounds whose constituents have the relationship of "equalness", namely the coordinative compounds.

6.2.3 Coordinative (copulative) compounds

According to Stevanović (1964: 425) coordinative (also 'copulative' or 'parallel') compounds are defined in the following way: "naporedne ili kako inače, iako im kopula često nedostaje, ne bez razloga, zovu kopulativne jesu složenice čije delove u jedno celinu vezala istovetnost njihove funkcije [parallel, or commonly known as copulative compounds because they miss the copula, are those compounds whose both elements have been unified into one entity by their identical function]". Further on he adds that this function can be one of the following: subject, attribute, predicate or object, as in *sjeveroistok* ('northeast'), *jugoistok* ('southeast'), *gluvonijem* ('deaf and dumb'), etc. When it comes to the term 'copulative compounds', Klajn (2002: 33) points out that "umesto termina *kopulativne*, koji bi se mogao shvatiti kao da upućuje na obavezno prisustvo nekog trećeg elementa – kopule [...], verovatno je bolje govoriti o *naporednim* [...] ili *koordinativnim* složenicama [instead of the term copulative compounds that could refer to the obligatory existence of some third element – copula, it is probably more appropriate to talk about parallel or coordinative compounds]". Apart from this, Stevanović's and Klajn's descriptions do not match in another point. Namely, Stevanović (426) claims that those compounds are either nominal or adjectival and both of its elements are originally either nouns or adjectives and that that their previous function does not allow these two word classes to be joined within one copulative compound. However, Klajn (33) refutes this claim by giving examples of compounds that consist of two adverbs (*katkad* ('sometimes')), two numbers (*pet-šest* ('five or six')), or two pronouns (*štošta* ('something')). He also adds these examples can be seen as half-compounds. However, as the topic of this thesis is nominal compounds, this type will not be discussed in greater detail. In contrast to the determinative and rectional compounds that can be analysed according to the

formula $AB=B$, copulative compounds can be represented in the following way: $AB=B$, as well as $AB=A$. Thus, it can be assumed that these compounds have two semantic heads, since *sjeveroistok* ('north east') is both, north and east. Further on, Rammelmeyer (1975: 48) states that

[b]ei diesen [Kopulativkomposita] ergibt sich die Gesamtsemantik des Kompositums aus der Addition der Kompositionsglieder; ein determinatives Verhältnis zwischen Vorder- und Hinterglied besteht nicht; die Folge der Kompositionsglieder wäre daher im Grunde ohne semantische Veränderung des Kompositums umkehrbar, wäre sie nicht durch Konvention festgelegt.

As already stated, when the morphological pattern of this type of compounds is concerned, the most productive types are noun + noun and adjective + adjective. Therefore let us start with the coordinative compounds that are composed of two nouns. But before doing any further analysis, it should be remarked that these compounds are not typical of Serbian, nor of other Slavic languages. Generally speaking, "[d]ie Gruppe der Kopulativkomposita ist im Serbokroatischen ebenso wie in den anderen slavischen und auch westeuropäischen Sprachen wenig produktiv und zahlenmäßig begrenzt." (Rammelmeyer 1975: 48) However, some examples can be found. For instance, *bogočovjek* ('God the Son') from *bog* ('the God') + *o* + *čovjek* ('man') is someone who is at the same time the god and the son (Jesus). Apart from this possibility, the compound can be analysed as 'božanski čovjek' ('the divine man'). In this context, it is interesting to mention compounds that seemingly have the same structure as *bogočovjek*. Namely, compounds such as *bogomajka* ('the mother of the God') from *bog* ('the God') + *o* + *majka* ('mother') and *bogosestra* ('the Sister of the God') from *bog* ('the God') + *o* + *sestra* ('sister') are without any doubt determinative compounds as they cannot be analysed as persons who are at the same time 'the mother and the god' and 'the sister and the god' respectively, but rather as *mother* and *sister* whose meaning has been narrowed down by the first element. According to Rammelmeyer (49) there is one more compound that could be analysed in the same way as *bogočovjek*, namely *majmunočovjek* ('ape man') from *majmun* ('ape') + *o* + *čovjek* ('man'). However, he also points out that it is possible to derive the attributive meaning from the first element, and thus the compound could be analysed as 'an ape-like person'. The compound *strahopoštovanje* ('awe, veneration') from *strah* ('fear') + *o* + *poštovanje* ('respect'), is the translation of the

German compound *Ehrfurcht*. The element *poštovanje* could not become the first part of the compound probably because it is either too long or because it is a noun derived from a verb. Apart from the geographical terms that have been mentioned (*sjeveroistok*, *sjeverozapad*, *jugoistok*, *jugozapad*) typical of this group are also chemical terms, such as *hlorovodonik* ('hydrogen chloride'), *sumporvodonik* ('hydrogen sulphide'), etc. Additionally, the binding vowel *o* is characteristic for many other coordinative compounds that have foreign origins: *metaloplastika* ('metal and plastic'), *labiodental* ('labiodental'), *labiovelar* ('labiovelar'), *kumulonimbus* ('cumulonimbus'), etc.

6.3 Exocentric compounds

Until now compounds have been discussed according to the syntactic relationships between the elements of compounds. Thus, it has been stated that there are three types of compounds: determinative, rectional and copulative compounds. It has also been pointed out that the borderline between determinative and rectional compounds is very often unclear and thus it might be argued that they actually belong to the same group. When it comes to the semantic classification of compounds, it is important to take into account what Vukićević (1995: 136-138) calls "semantička usmerenost složenice [the semantic direction of the compound]". According to her, compounds are divided into endocentric and exocentric ones. Endocentric are those "kod kojih se značenje sastavnih delova pre srastanja u složenicu nije izmenilo unutar složenice [compounds whose elements retained the meaning they had before they were joined in a compound]" and exocentric compounds are those whose elements slightly changed their meaning. However, Klajn (2002: 37) argues that this definition cannot be accepted, since elements in all compounds somehow change or modify their meaning. For instance, *brodogradnja* ('shipbuilding') does not deal only with building ships, but also with the repair and the maintenance of ships. Therefore, the same definition that has been applied to English compounds must be applied to Serbian as well. That means that in endocentric compounds the meaning of the whole compounds is derived from the right-hand element if it is a determinative compound, i.e. *parobrod* ('steamboat') is a kind of ship; if it is a coordinative compound, each of the elements carries half of the meaning, i.e. *gluvonijem* ('deaf and dumb') is someone who is deaf as well as dumb. In

exocentric compounds, it is not possible to derive the meaning from either of its elements, and thus the meaning has to be reconstructed on the basis of the meanings of those two constituents.

It also has to be pointed out that when it comes to the classification of compounds into endocentric and exocentric ones, the only compounds that have been taken into account in the Serbian literature are noun compounds. Thus Vukićević (1994: 149) says that “imeničke složenice su uglavnom eksocentrične, za razliku od pridevskih složenica koje su uglavnom endocetrične [nominal compounds are predominantly exocentric, while adjectival compounds are mostly endocentric]”, however she neither deals with the adjectival compounds nor does she give any example of it. Klajn (2002: 38) argues that it might be impossible to find any adjectival compound that is an exocentric one. However, this paper does not deal with adjectival compounds and thus this problem will not be discussed any further. Let us come back to the exocentric noun compounds. As Klajn states “činjenica je da se o endo- i eksocentričnim složenicama ne govori mnogo u svetskoj lingvistici [it is a fact that one cannot find enough information about exocentric compounds in world literature]” and therefore when it comes to exocentric compounds in Serbian it is clear that a satisfying description can follow only after a systematic classification has been done. Nonetheless, an attempt to classify them has to be made. Some verbs in Serbian have a strong tendency to create compounds of this type, so that it is possible to talk about small lexical families with the same second part. When it comes to the morphological pattern of this type of compounds, it is noun base + verb base + Ø, and the verbs that build these word families are: *voditi* (‘lead’) (*vodovod* (‘water pipeline’), *naftovod* (‘oil pipeline’), *parovod* (‘steam pipeline’), *gasovod* (‘gas pipeline’), *cjevovod* (‘pipeline’), etc.), *mjeriti* (‘measure’) (*kišomjer* (‘pluviometer’), *toplomjer* (‘thermometer’), *strujomjer* (‘electricity meter’), *visinomjer* (‘altimeter’), *svjetlomjer* (‘light meter’), etc.), *jesti* (‘eat’) (*mesojed* (‘meat eater’), *biljojed* (‘plant eater’), *travojed* (‘grass eater’), *mravojed* (‘ant eater’), etc.), *rezati* (‘cut’) (*bakrorez* (‘engraving’), *drvorez* (‘woodcut’), *kamenorez* (‘stonecut’), *linorez* (‘linocut’), etc.), *braniti* (‘guard, defend’) (*blatobran* (‘mudguard’), *kišobran* (‘umbrella’), *suncobran* (‘parasol’), *gromobran* (‘lightning conductor’), etc.), *pisati* (‘write’) (*rukopis* (‘handwriting’), *šakopis* (‘handwriting’), *svrakopis* (‘bad

handwriting')⁴², *ljetopis* ('annual'), *časopis* ('magazine'), etc.), *tvoriti* ('make, create') (*dobrotvor* ('benefactor'), *zlotvor* ('malefactor'), *umotvor* ('creation'), etc.), *žderati* ('devour') (*mesožder* ('meat devourer'), *ljudožder* ('human devourer'), *zrnožder* ('seed devourer'), *travožder* ('grass devourer'), etc.), *metnuti* ('put') (*vatromet* ('fireworks'), *nogomet* ('football'), *rukomet* ('handball'), etc.), etc. When it comes to the verb *lomiti* ('break'), Klajn (2002: 55) argues that there are only four compounds that have this verb as its second element and they have completely different semantic structures: *kamenolom* ('quarry'), *brodolom* ('shipwreck'), *vratolom* ('a clumsy person') and *kostolom* ('bone fracture').

As this group includes an extremely large number of compounds, it is to be expected that the semantics of these words is rather heterogeneous. Many of them denote devices, gadgets (e.g. *naftovod*, *toplomjer*, *gromobran*), some of them are used to represent states and events (*zemljotres*, *vatromet*), results of an action (*bakrorez*, *rukopis*), devices for performing certain actions (*mišomor* ('rat poison')), human beings (*bogolslov* ('theologian'), *rodoljub* ('patriot')), plants (*suncokret*) (sunflower), etc. In most of the cases the noun represents the object of the verb, as in *mesožder* ('ždere meso'), subject, as in *vodopad* ('voda pada'), an instrumental complement, as in *rukomet* ('metnuti rukom'), etc.

⁴² This compound denotes handwriting that is illegible or extremely difficult to read.

7. Conclusion

The aim of this thesis has been to describe and compare noun-noun compounds in English and Serbian. This class of compounds is known as the biggest and the most productive one in both languages. Even though these two languages are typologically different, it has been seen that there are many similarities between them, especially when it comes to the categorization of compounds. However, before it was possible to analyse compounds and their categories, it was necessary to discuss the problem that arose in connection with the definition and the status of the terms 'compound' and 'compounding'. Further more, the most important features of compounds were analysed and an attempt was made to settle the problem with the definition of compounds.

Firstly, it has to be stated that compounds in both languages are complex lexemes, which means that they are built of some smaller units, and these smaller units, or constituents are in a certain relationship and this relationship is the starting point for the classification of compounds. As has been seen, the classification of compounds largely depends on the aspects one looks at. Thus, when it comes to the syntactic properties of noun-noun compounds, it has been stated that both languages classify compounds into subordinative and coordinative ones. It has also been shown that coordinative compounds in Serbian are divided into two subgroups, determinative and relational compounds. However, Serbian linguists tend to classify them as one group as the differences between relational and determinative compounds are often so blurred that it is not possible to say to which subgroup a certain compound belongs. Coordinative (also copulative) compounds in both languages have been defined in a similar way. Namely, they are defined as compounds in which a copula can be inserted between constituents. Later, however, it has become clear quickly that only the syntactic aspect is not enough and therefore it was necessary to take semantic features into account, thus bringing us a new class of compounds, namely exocentric ones. However, as the number of exocentric noun-noun compounds in English, as well as in Serbian, is very small, other combinations, such as adjective-noun, verb-noun or verb-preposition, have had to be taken into account, otherwise the analysis of type would not have been possible. Finally, it has also become obvious that some pragmatic aspects have had to be taken into account in the presentation

of the major categories. Further on it has also been stated that our general knowledge of the world plays a very important role in decoding the meaning of compounds, since neither semantic nor syntactic relationships between constituents have been enough to clarify the intended meaning of a compound.

As has been observed, there is a number of different types of compounds in both languages and it can be stated that both languages are very rich in this type of compounds, although one could observe that compounds in general are a more typical phenomenon in English, than in Serbian. Whatever the case may be, it is a fact that this is the most productive way of producing new words in English and also a very productive one in Serbian and thus it is to be expected that new compounds constantly enter both languages.

8. References

- Adams, Valerie. 1973. *An introduction to modern English word formation*. (English language series; 7). London: Longman.
- Adams, Valerie. 2001. *Complex words in English*. Harlow: Pearson Education.
- Babić, Stjepan. 1986. *Tvorba riječi u hrvatskom književnom jeziku: nacrt za gramatiku*. [Word-formation in Croatian: a sketch for the grammar]. Zagreb: Djela Jugoslovenske Akademije Znanosti i Umjetnosti.
- Barić, Eugenija. 1980. *Imeničke složenice neprefiksarne i nesufiksarne tvorbe* [Unprefixed and unsuffixed noun compounds]. (Znanstvena biblioteka hrvatskog filološkog društva; 11). Zagreb: Liber.
- Belić, Aleksandar. 1949. *Savremeni srpskohrvatski književni jezik. II deo: nauka o gradjenju riječi*. [The contemporary Serbo-Croatian literary language. Part II: word-formation]. Beograd: Preduzeće narodne republike Srbije.
- Bauer, Laurie. 1978. *The grammar of nominal compounding: with special reference to Danish, English and French*. Odense: Odense University Press.
- Bauer, Laurie. 1979. "On the need for pragmatics in the study of nominal compounding". *Journal of pragmatics* 3, 45-50.
- Bauer, Laurie. 1983. *English word-formation*. (Cambridge textbooks in linguistics). Cambridge: Cambridge University Press.
- Bauer, Laurie. 1998. "When is a sequence of two nouns a compound in English?" *English language and linguistics* 2, 65-86.
- Bauer, Laurie. 2001. "A corpus-based study of compounding in English". *Journal of English linguistics* 29, 101-123.
- Bauer, Laurie. 2003. *Introducing linguistic morphology*. (2nd edition). Edinburgh: Edinburgh University Press.
- Benczes, Réka. 2006. *Creative compounding in English: the semantics of metaphorical and metonymical noun-noun combinations*. (Human cognitive processing; 19). Amsterdam: Benjamins.
- Brekle, Herbert Ernst. 1976. *Generative Satzsemantik im System der englischen Nominalkomposition*. München: Wilhelm Fink Verlag.
- Carstairs-McCarthy, Andrew. 2002. *An introduction to English morphology: words and their structure*. (Edinburgh textbooks on the English language). Edinburgh: Edinburgh University Press.

- Coseriu, Eugenio. 1977. "Inhaltliche Wortbildungslehre (am Beispiel des Typs 'coupe-papier')." In Brekle, Herbert E.; Kastovsky, Dieter. (eds.). *Perspektiven der Wortbildungsforschung: Beiträge zu Wuppertaler Wortbildungskolloquium vom 9.-10. Juli 1976*. Bonn: Bouvier, 48-61.
- Dalton-Puffer, C.; Plag, I. 2001. "Categorywise, some compound-type morphemes seem to be rather suffix-like: on the status of *-ful*, *-type*, and *-wise* in Present Day English". *Folia Linguistica* 34, 225-244.
- Downing, Pamela. 1977. "On the creation and use of English compound nouns." In Katamba, Francis (ed.). *Morphology: critical concepts in linguistics*. London and New York: Routledge, 169-208.
- Fernández-Domínguez, Jesús. 2008. "Productivity measurement of English compounding based on a corpus of the nominal type". PhD thesis, University of Jaén, Jaén.
- Gagné, Christina L.; Spalding, Thomas L. 2010. "Relational competition during compounding interpretation". In Scalise, Sergio; Vogel, Irene (eds.). *Cross-disciplinary issues in compounding*. (Amsterdam studies in the theory and history of linguistic science, series 4: Current issues in linguistic theory; 311). Amsterdam: Benjamins, 287-300.
- Gimborn, Germeid. 2001. "Bahuvrīhi compounds in Old English: a comparative study of exocentric compounds in Germanic". Diplomarbeit, Universität Wien, Wien.
- Jespersen, Otto. 1942. *A modern English grammar on historical principles, Part VI: Morphology*. Copenhagen: Ejnar Munksgaard.
- Kapetanović, Amir. 2007. "Jesu li endocentrične imeničke složenice tvorbeni inovacija u hrvatskom jeziku 19. stoljeća? [Are the endocentric nominal compounds a word-formation innovation in the Croatian language in the 19th century?]" *Rasprave instituta za hrvatski jezik i jezikoslovlje* 33, 235-243.
- Kastovsky, Dieter. 1982. *Wortbildung und Semantik*. (Studienreihe Englisch 14). Düsseldorf: Schwann-Bagel.
- Kastovsky, Dieter. 1986. "Word-formation and pragmatics". In Rauchbauer, Otto (ed.). *A yearbook of studies in English language and literature 1985/86*:

Festschrift für Siegfried Korningen. (Wiener Beiträge zur Englischen Philologie 80). Wien: Wilhelm Braumüller, 63-78.

Kastovsky, Dieter. 2000. "Words and word-formation: morphology in OED". In Mugglestone, Lynda (ed.). *Lexicography and the OED: pioneers in the untrodden forest*. Oxford: Oxford University Press, 110-125.

Kastovsky, Dieter. 2008. "Special linguistic course: morphology II – English word-formation". Lecture held at the Department of English, University of Vienna, October 2008 – January 2009.

Klajn, Ivan. 2002. *Tvorba reči u savremenom srpskom jeziku, Kijiga I: slaganje i prefiksacija*. (Prilozi gramatici srpskoga jezika 1). [*Word-formation in contemporary Serbian, Book I: compounding and prefixation*]. Beograd: Zavod za udžbenike i nastavna sredstva.

Klajn, Ivan. 2003. *Tvorba reči u savremenom srpskom jeziku, Kijiga II: sufiksacija i konverzija*. (Prilozi gramatici srpskoga jezika 2). [*Word-formation in contemporary Serbian, Book II: suffixation and conversion*]. Beograd: Zavod za udžbenike i nastavna sredstva.

Komkov, Stephanie. 1983. "The compounds substantives of Serbo-Croatian and Russian, with a sketch of the compound substantives of common Slavic". PhD thesis, Yale University.

Lang, Kathrin. 2007. "Nominal compounds in English and French: a contrastive analysis". Diplomarbeit, Universität Wien, Wien.

Levi, Judith N. 1978. *The syntax and semantics of complex nominals*. New York: Academic Press.

Lieber, Rochelle. 1983. "Argument linking and compounds in English." In Katamba, Francis (ed.). *Morphology: critical concepts in linguistics*. London and New York: Routledge, 209-247.

Lieber, Rochelle. 2005. "English word-formation processes: observations, issues, and thoughts on future research". In Štekauer, Pavol; Lieber, Rochelle (eds.). *Handbook of word-formation. (Studies in inatural language and linguistic theory 64)*. Dordrecht, Springer, 375-427.

Lipka, Leonhard. 1972. *Semantic structure and word-formation : verb-particle constructions in contemporary English*. (Internationale Bibliothek für allgemeine Linguistik; 17). München: Fink.

- Lipka, Leonhard. 1990. *An outline of English lexicology: lexical structure, word semantics, and word-formation*. (Forschung & Studium Anglistik; 3). Tübingen: Max Niemeyer Verlag.
- Marchand, Hans. 1969. *The categories and types of present day English word-formation: a synchronic-diachronic approach*. (2nd edition). München: Beck.
- Marchand, Hans. 1974a. "Synchronic analysis and word-formation." In Kastovsky, Dieter (ed.). *Studies in syntax and word-formation: selected articles*. (International library of general linguistics 18). München: Fink, 171-184.
- Marchand, Hans. 1974b. "Die Länge englischer Komposita und die entsprechenden Verhältnisse im Deutschen." In Kastovsky, Dieter (ed.). *Studies in syntax and word-formation: selected articles*. (International library of general linguistics 18). München: Fink, 199-205.
- Marchand, Hans. 1974c. "The analysis of verbal nexus substantives." In Kastovsky, Dieter (ed.). *Studies in syntax and word-formation: selected articles*. (International library of general linguistics 18). München: Fink, 276-291.
- Marchand, Hans. 1974d. "On the analysis of substantive compounds and suffixal derivatives not containing a verbal element". In Kastovsky, Dieter (ed.). *Studies in syntax and word-formation: selected articles*. (International library of general linguistics 18). München: Fink, 292-322.
- Marchand, Hans. 1974e. "Expansion, transposition, and derivation." In Kastovsky, Dieter (ed.). *Studies in syntax and word-formation: selected articles*. (International library of general linguistics 18). München: Fink, 322-337.
- Maretić, Tomo. 1899. *Gramatika i stilistika hrvatskoga ili srpskoga književnog jezika*. [The grammar and stylistics of the Croatian or Serbian literary language]. Zagreb:
- Matthews, Peter H. 1974. *Morphology: an introduction to the theory of word-structure*. (Cambridge textbooks in linguistics). Cambridge: Cambridge University Press.
- Olsen, Susan. 2000a. "Composition". In Booij, Geert; Lehmann, Christian; Mugdan, Joachim (eds.). *Morphology: an international handbook on inflection and word-formation*. Berlin: Walter de Gruyter, 897-916.
- Olsen, Susan. 2000b. "Compounding and stress in English: a closer look at the boundary between morphology and syntax". *Linguistische Berichte* 181, 55-69.
- Olsen, Susan. 2000c. "Copulative compounds: a closer look at the interface between syntax and morphology". In Booij, Geert; von Marle, Jaap (eds.). *Yearbook of morphology*. Dordrecht: Kluwer, 279-320.

- Oshita, Hiroyuki. 1994. "Compounds: a view from suffixation and a-structure alteration". In Booij, Geert; von Marle, Jaap (eds.). *Yearbook of morphology*. Dordrecht: Kluwer, 179-205.
- Perkles, Maria. 2008. "Compounds in English and Russian: a comparative analysis". Diplomarbeit, Universität Wien, Wien.
- Plag, Ingo. 2003. *Word-formation in English*. Cambridge: Cambridge University Press.
- Rakić, Stanimir. 1998. "O nominalizaciji i nasleđivanju strukture argumenata [About nominalization and the inheritance of the argument structure]". *Južnoslovenski filolog* 54, 15-33.
- Rakić, Stanimir. 2001. "O morfološkom pojmu glave, konverziji u strukturi složenica [About the morphological term 'head', conversion and the structure of compounds]". *Južnoslovenski filolog* 57, 7-31.
- Rakić, Stanimir. 2007. "O leksičkom sloju imeničkih složenica tipa imenica + imenica u srpskom jeziku [About the lexical level of the compound nouns of the type noun + noun in the Serbian language]". *Zbornik Matice srpske za filologiju i lingvistiku* 50(1-2), 767-792.
- Rufener, John. 1971. *Studies in the motivation of English and German compounds*. Zurich: Juris.
- Ryder, Mary Ellen. 1994. *Ordered chaos: the interpretation of English noun-noun compounds*. (University of California publications in linguistics; 127). Berkley: University of California Press.
- Schaefer, Susanna. 2000. "A contrastive analysis of English and German nominal compounds". Diplomarbeit, Universität Wien, Wien.
- Selkirk, O. Elisabeth. 1982. *The syntax of words*. (Linguistic Inquiry Monographs; 7). Cambridge, Massachusetts: The MIT Press.
- Shaw, James Howard. 1979. *Motivierte Komposita in der deutschen und englischen Gegenwartssprache*. (Tübinger Beiträge zur Linguistik; 118). Tübingen: Narr.
- Stanojčić, Živojin; Popović, Ljubomir. 2004. *Gramatika srpskoga jezika: udžbenik za I, II, III i IV razred srednje škole. [The grammar of the Serbian language: the book for I, II, III and IV grade of the grammar school]*. (9th edition). Beograd: Zavod za Udžbenike i Nastavna Sredstva.
- Štekauer, Pavol. 1998. *An onomasiological theory of English word-formation*. (Studies in functional and structural linguistics; 46). Amsterdam: Benjamins.
- Štekauer, Pavol. 2000. *English word-formation: a history of research (1960 - 1995)*. Tübingen: Narr.

- Štekauer, Pavol. 2005. *Meaning predictability in word formation: novel, context-free naming units*. (Studies in functional and structural linguistics; 54). Amsterdam: Benjamins.
- Stevanović, Mihailo. 1975. *Savremeni srpskohrvatski jezik, Knjiga I: uvod, fonetika, morfologija*. [*The contemporary Serbo-Croatian language, Book I: introduction, phonetics, morphology*]. (3rd edition). Beograd: Naučna knjiga.
- Vukićević, Dušanka. 1994. "O značenju imeničkih složenica [About the meaning of noun compounds]". *Zbornik Matice srpske za filologiju i lingvistiku*, 37, 145-149.
- Vukićević, Dušanka. 1995. "Imeničke složenice u savremenom srpskom književnom jeziku [Noun compounds in the contemporary literary Serbian language]". *Zbornik Matice srpske za filologiju i lingvistiku*, 38/1, 127-174.
- Warren, Beatrice. *Semantic pattern of noun-noun compounds*. Götteborg: Acta Universitatis Gothoburgensis.
- Yule, George. 1996. *The study of language*. (2nd edition). Cambridge: Cambridge University Press.

9. Appendix

9.1 German summary

Das Ziel dieser Arbeit ist eine ausführliche und umfassende Beschreibung der Komposita, die aus zwei Substantiven bestehen, im Serbischen und Englischen darzustellen. Diese Komposita werden auch Nominal-Komposita genannt. Durch die Beschreibung werden die wichtigsten Merkmale dieser Gruppe zuerst beschrieben und dann miteinander verglichen. Komposita gehören zur Gruppe der so genannten komplexen Wörter, das heißt, sie bestehen aus zwei oder mehreren Morphemen. In beiden Sprachen werden Komposita aus schon existierenden Morphemen gebildet. Dabei wird gezeigt, dass alle Komposita motiviert sind, also aus Elementen bestehen, die eine Bedeutung haben. Allerdings ist die Art und Weise, wie diese Morphemen miteinander verknüpft werden, in beiden Sprachen unterschiedlich. So werden zum Beispiel die Nominal-Komposita im Englischen derart gebildet, dass zwei oder mehrere Elemente verknüpft werden, während im Serbischen zwischen den Elementen meistens ein Bindevokal eingesetzt wird. Auch wenn die Komposita aus mehr als zwei Morphemen bestehen, können sie immer auf eine binäre Beziehung reduziert werden.

Bevor die einzelnen Komposita-Kategorien behandelt werden, ist es allerdings notwendig, die Frage zu beantworten, wie ein „Wort“ zu definieren ist. Dabei gibt es mehrere Gesichtspunkte, die in Betracht genommen werden müssen, wie etwa die Schreibweise, die phonologischen sowie die semantischen und syntaktischen Gesichtspunkte. Nachdem die Definition des Terminus „Wort“ festgelegt ist, ist es nun möglich, die wichtigsten allgemeinen Merkmale der Nominal-Komposita in den beiden Sprachen anhand der folgenden Kriterien darzustellen: die Betonung, die Funktion des Kopfes eines Kompositums sowie die morphologische und die syntaktisch-semantische Struktur des Kompositums. In der Beschreibung der serbischen Komposita wird auch der diachronische Aspekt der Wortbildung berücksichtigt, da etliche Komposita nur so erklärt werden können.

Der zweite Teil der Arbeit umfasst eine detaillierte Übersicht der verschiedenen Gruppen der Nominal-Komposita in den beiden Sprachen. Die Komposita im Englischen werden aufgrund zweier Kriterien – die syntaktischen und die semantischen – unterteilt. Bei der syntaktischen Betrachtung erfolgt die

Unterteilung je nach dem, ob sich das linke und rechte Element des Kompositums in einer subordinativen oder koordinativen Relation befinden. Im Falle von subordinativen Relationen spricht man von Expansionen, die wiederum in primäre und sekundäre Komposita untergliedert werden. Im Falle von koordinativen Relationen sind diese Komposita als so genannte kopulative Komposita zu betrachten. In beiden Fällen befindet sich der semantische Kopf innerhalb des Kompositums, weshalb sie auch als endozentrische Komposita bezeichnet werden. Im Gegensatz dazu werden Komposita, bei denen der außerhalb des Kompositums steht, als exozentrische Komposita bezeichnet.

Die Unterteilung der Komposita im Serbischen erfolgt aufgrund der selben (syntaktischen und semantischen) Kriterien mit Ausnahme der Untergruppen der Expansionen, die im Serbischen als determinative Komposita bezeichnet werden. Jedoch muss darauf hingewiesen werden, dass manche LinguistInnen zwischen determinativen und rektionalen Komposita unterscheiden. Allerdings wird diese Theorie von den meisten LinguistInnen abgelehnt, weshalb alle subordinative Komposita als eine Gruppe betrachtet werden. Ebenfalls gibt es – wie im Englischen – die kopulativen und exozentrischen Komposita. Wie oben bereits erwähnt ist in den meisten serbischen Komposita ein Bindevokal zu finden, der aber bedeutungslos ist und meistens nur dazu dient, die Aussprache der Wörter zu erleichtern. Darüber hinaus gibt es in beiden Sprachen Komposita, die in keine der beschriebenen Gruppen zugeordnet werden und auch sonst nicht erklärt werden können.

9.2 Curriculum vitae

Persönliche Daten

Geburtsdatum:	1. März 1985
Geburtsort:	Olten, Schweiz
Staatsbürgerschaft:	Bosnien und Herzegowina

Ausbildung

Juni 2004	Matura am allgemeinen Gymnasium in Bijeljina, Bosnien und Herzegowina
Okt. 2004 – Jän. 2005	Besuch eines Deutschkurses mit erfolgreichem Abschluss
März 2005	Beginn des Studiums Anglistik und Amerikanistik an der Universität Wien
Okt. 2006	Beginn des Studiums Skandinavistik im Rahmen der freien Wahlfächer
Aug. 2007	Besuch eines dreiwöchigen Dänischkurses an der Internationalen Hochschule in Helsingør, Dänemark
Jän. 2008	Abschluss des ersten Studienabschnitts des Studiums Anglistik und Amerikanistik
Okt. 2008 – Jän. 2009	Auslandssemester am Institut für Englisch und am Institut für Nordische Studien and der Universität Kopenhagen
März 2010	Beginn des Bachelor-Studiums der Skandinavistik an der Universität Wien